

NBER WORKING PAPER SERIES

EFFICIENCY VERSUS EQUITY IN THE ALLOCATION OF MEDICAL SPECIALTY  
TRAINING POSITIONS IN SPAIN:  
A HEALTH POLICY SIMULATION BASED ON A DISCRETE CHOICE MODEL

Jeffrey E. Harris  
Beatriz G. López-Valcárcel  
Patricia Barber  
Vicente Ortún

Working Paper 19896  
<http://www.nber.org/papers/w19896>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
February 2014

This project was funded by the Spanish Ministry of Science and Innovation through Grant ECO2010–21558 to BGLV as the principal investigator. The funder had no influence in the conduct of this study or the drafting of this manuscript. We gratefully acknowledge the Ministry of Health, Social Services and Equality, General Sub-Directorate on Professional Planning, Spain, for providing us the 2012 MIR registry data. We acknowledge the many medical students and MIR candidates who responded to our 2011 and 2012 surveys. We thank Jaime Pinilla Domínguez for technical assistance. Finally, we thank the participants in workshops and seminars that we have given concerning this research, including: the Pontificia Universidad Católica, Chile (JEH), Universidad Pública de Navarra (BGLV), Universidad de Vigo (BGLV), and the Grupo Evaluación de Políticas y Servicios de Salud, Asociación de Economía de la Salud (EvaluAES), Valencia (VO). The opinions expressed in this paper are ours and ours alone, and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Jeffrey E. Harris, Beatriz G. López-Valcárcel, Patricia Barber, and Vicente Ortún. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Efficiency versus Equity in the Allocation of Medical Specialty Training Positions in Spain: A Health Policy Simulation Based on a Discrete Choice Model

Jeffrey E. Harris, Beatriz G. López-Valcárcel, Patricia Barber, Vicente Ortún

NBER Working Paper No. 19896

February 2014

JEL No. C25,C78,D82,I11,I18

**ABSTRACT**

**Background.** In Spain’s “MIR” system of allocating residency training positions, medical school graduates are ranked according to their performance on a national exam and then sequentially choose from the remaining available training slots. We studied how changes in the MIR system might address the inadequate supply of practitioners of family and community medicine in that country.

**Data.** Our data included: a registry of the actual residency positions chosen by medical school graduates in the 2012 MIR cycle; a 2012 post-MIR survey in which graduates made counterfactual choices as to what they would have chosen but for their position in the national rankings; and a 2011 survey of the relative importance of specialty attributes among final-year medical students in the same cohort.

**Methods.** We modeled the MIR system as a one-sided matching mechanism based priority rankings, also called “serial dictatorship.” Within this model, we developed a framework for evaluating the tradeoff between the efficiency gains from increasing the supply of practitioners of family and community medicine and the equity-related benefits of permitting the most talented medical students to make their specialty choices first. We then applied our framework to real data on medical school graduates’ specialty choices during 2012 MIR cycle. Our empirical analysis, based on the multinomial logit model with random coefficients, took account of the endogeneity of choice sets induced by the MIR scheme. We then used the parameter estimates to simulate various alternative public policies, including random ranking of candidates, restrictions on the supply of training positions, and policies designed to upgrade medical school graduates’ valuations of a career in family and community medicine.

**Results:** Both random ranking and restrictions in supply resulted in a relatively small efficiency gains from training more productive medical school graduates in family and community medicine, but at the same time a substantial equity losses. Improvements in two key attributes of family and community medicine – professional prestige and the proportion of income from private practice – resulted in substantial gains in both equity and efficiency.

**Conclusions:** Policies designed to increase the prestige and remuneration of practitioners of family and community medicine have the potential to be more efficient and equitable than other alternatives.

Jeffrey E. Harris  
Department of Economics, E18-258  
Massachusetts Institute of Technology  
Cambridge, MA 02139  
and NBER  
jeffrey@mit.edu

Beatriz G. López-Valcárcel  
Department of Quantitative Methods  
in Economics and Management  
University of Las Palmas de Gran Canaria  
Campus Tafira  
Las Palmas, Gran Canaria, Spain  
bvalcarcel@dmc.ulpgc.es

Patricia Barber  
Department of Quantitative Methods  
in Economics and Management  
University of Las Palmas de Gran Canaria  
Campus Tafira  
Las Palmas, Gran Canaria, Spain  
pbarber@dmc.ulpgc.es

Vicente Ortún  
Faculty of Economic and Business Sciences  
Universitat Pompeu Fabra  
Barcelona, Spain  
vicente.ortun@upf.edu

## 1. Introduction

Numerous authors in many countries have lamented the shortage of primary care physicians, particularly the inadequate supply of practitioners of family and community medicine (Bodenheimer 2006, Rosenblatt, Andrilla et al. 2006, Scott, Larson et al. 2006, Mariolis, Mihas et al. 2007, Colwill, Cultice et al. 2008, Thistlethwaite, Leeder et al. 2008, Huibers, Giesen et al. 2009, Steinbrook 2009, Barber and Lopez-Valcarcel 2010). The roots of the shortage are complex and varied, but undoubtedly include noncompetitive distortions in the relative prices of primary, secondary and tertiary medical care, as well as undervalued external benefits of preventive and other community-based services (Watts and Segal 2009, Cheng, Scott et al. 2011, de Bakker, Struijs et al. 2012). Whatever the underlying causes, there has been a general perception that the social value of training an additional primary care doctor exceeds the private gains, and that the market has failed to take this divergence between private and social benefits into account.

A wide variety of corrective measures have been proposed, including changes in payment mechanisms and physician compensation, improvements in working conditions, policies to counter the low prestige of primary care medicine, and the training of non-physician practitioners as substitutes (Dorsey, Jarjoura et al. 2003, Thornton and Esposto 2003, Gagne and Leger 2005, Goroll, Berenson et al. 2007, Ortun, Gonzalez Lopez-Valcarcel et al. 2008, Krueger and Halperin 2010, Sivey, Scott et al. 2010, Gonzalez Lopez-Valcarcel and Barber Perez 2012). However, there has been relatively little consideration of alternative policies focused at the critical point where medical school graduates choose residency training positions. A growing body of theoretical research on matching mechanisms ultimately led to the reengineering of the U.S. residency matching program (Roth and Peranson 1999). Still, the principal focus of such research – and its subsequent applications to public school assignment, housing, and other matching problems (Abdulkadiroğlu and Sönmez 1998, Ergin 2002, Abdulkadiroğlu and Sönmez 2003, Pathak 2011) – has been the development of algorithms that result in stable, Pareto-efficient allocations. The design of matching mechanisms that correct for externalities and other market failures has not been a priority.

Our objective in this article is to evaluate policy alternatives to the current national system for allocating residency training positions in Spain, widely known as “MIR.” We develop a simple theoretical model of the MIR system as a one-sided matching mechanism based on

priority rankings. The model incorporates the possibility of an external benefit from training a doctor in family and community medicine. It also explicitly takes into account the tradeoff between efficiency – in the sense of correcting for the undervaluing of primary care physicians – and equity – in the vertical sense that the most talented medical students deserve to make their specialty choices first. Within this theoretical model, we explore the properties of various modifications of the MIR scheme. We then apply our conceptual models to real data on MIR outcomes, and use our empirical estimates to simulate the effects of alternative policy modifications of the MIR system.

In the following section, we describe the institutional details of the MIR system of allocating residency training positions. Section 3 presents our theoretical framework for evaluating the tradeoff between equity and efficiency within the MIR context. In Section 4, we present our data: a registry of the actual choices of medical school graduates in the 2012 MIR cycle; a survey in which graduates made counterfactual choices as to what they would have chosen but for their position in the queue; and a survey of the perceived attributes of each specialty among final-year medical students carried out in 2011. Section 5 presents our model of discrete choice, a logit model with random coefficients and endogenous constraints on candidates' choice sets imposed by the MIR allocation system.

In Section 6, we present our strategy for simulating alternative policy scenarios. We evaluate four different policies: (1) random assignment of candidates' ranking in the priority mechanism; (2) concentrated reduction of supply, eliminating training positions in small centers; (3) generalized reduction of supply, reducing the number of positions in all centers proportionately; and (4) policies designed to selectively increase the prestige and remuneration of practitioners of family and community medicine.

Section 7 presents the results of our discrete choice models, while Section 8 shows the results of our policy simulations. We find that policies designed to increase the prestige and remuneration in private practice of practitioners of family and community medicine have the potential to be more efficient and equitable than other alternatives. Finally, in Section 9 we discuss the limitations of our work and offer a roadmap for further research.

## 2. The MIR System of Allocating Residency Training Positions in Spain

The allocation of residency training positions in Spain is organized and regulated at the national level by a system widely known as MIR, which stands for “médico interno residente,” literally “resident medical intern.” On an annual basis, the central government’s Ministry of Health authorizes postgraduate training programs in 47 specialties. The distribution of specialty training programs and the number of training positions within each program are based on models of population needs and physician supply (Barber and Gonzalez Lopez-Valcarcel 2010). To be considered for a residency training position during a particular annual cycle, each “candidate” must have an approved diploma from a Spanish or foreign medical school and take a national examination. Candidates are then ranked on the basis of their MIR combined score, which is a weighted average of their national exam score (90%) and medical school grade point average (10%).

Once all training positions are authorized and all candidates are nationally ranked, the final phase of the annual MIR cycle functions essentially as a one-sided sequential allocation mechanism or “serial dictatorship,” in which the training programs play only a passive role. The top-ranked candidate chooses his preferred residency training position from the entire set of nationally available training programs. Then the second-ranked candidate chooses from the remaining available residency positions, and the process continues iteratively until all training positions are exhausted or all candidates have elected positions.

The national rank ordering is a critical element of the MIR allocation scheme, and in a loose sense resembles a tournament where the prizes are highly valued residency training positions in the most sought-after specialties. With each annual MIR cycle, the graduates of the nation’s top medical schools consistently attain the highest combined scores and thus get to choose their prizes first (Gonzalez Lopez-Valcarcel, Ortun et al. 2013, Lopez-Valcarcel, Ortun et al. 2013). Figure 1 highlights the superior performance of the 213 graduates of the medical school of the Universidad Autónoma de Madrid, a university that has consistently produced top-ranked candidates, among a total of 6,348 candidates in the 2013 MIR cycle.

Year after year, the highest scoring candidates from the top universities consistently choose such specialties as plastic surgery, dermatology and cardiology. At the bottom of the national ranking, the residual claimants are left with a Hobson’s choice between enrolling in a residency in family and community medicine or dropping out in order to retake the national

exam the following year (Gonzalez Lopez-Valcarcel, Barber Perez et al. 2011). In the 2002 MIR cycle, as shown in Figure 2, only 200 candidates within the top-ranked 3,000 chose a residency position in family and community medicine, and by the 2012 cycle, only 50 had done so.

Figure 2 suggests that the declining status of family and community medicine may have reached bottom in 2010 and, in fact, there is evidence of a minor reversal (Harris, Gonzalez Lopez-Valcarcel et al. 2013). The critical factor appears to be job security. Until the financial crisis erupted in late 2008, unemployment among Spanish physicians was a relatively rare phenomenon, with most doctors obtaining long-term jobs as salaried employees of the Spanish national health system. As the budgetary cutbacks began to hit the health departments of the country's autonomous communities, an increasing number of physicians came every month to government unemployment offices looking for a job. In the face of rising unemployment, practitioners of family and community medicine experienced fewer job losses than specialists (Figure 3).

The prospects for increased job security in family and community medicine, however, may turn out to be short lived. Many recent graduates have been compelled to accept temporary employment contracts under conditions substantially inferior than those prevailing before the crisis. One sign of the increased turnover is the abrupt rise in the proportion of new employment contracts that go to practitioners of family and community medicine (Figure 4).

### 3. Theoretical Framework

We assume that there are  $N > 0$  candidates for residency training positions. We denote the set of candidates by  $C = \{c_i | i = 1, \dots, N\}$ . We further assume that there are  $M > 0$  available training positions, and denote the set of residency training positions by  $P = \{p_j | j = 1, \dots, M\}$ .

For simplicity, we assume that there are only two types of residency training programs: family medicine ( $F$ ), and a hospital-based specialty ( $H$ ). Our results could readily be generalized to more than two types of programs.<sup>1</sup> Let  $M_F$  and  $M_H$ , respectively, denote the number of training positions available in each of the two specialties, where  $M_F + M_H = M$ . We assume for

---

<sup>1</sup> Formally, the specialty is known as “family and community medicine” (or “*medicina familiar y comunitaria*”), but for brevity here we shall sometimes refer to it simply as “family medicine.”

now that the number of training positions equals or exceeds the number of candidates, that is,  $M \geq N$ . Violations of this assumption are discussed below. We further assume that candidates regard all positions within a specialty as interchangeable. The only characteristic of a residency training program that matters to a candidate is the specialty. We can thus regard the set  $P$  of training positions as having binary elements, equal to either 1 for family medicine or 0 for the hospital-based specialty.

We define an *assignment* as a single-valued mapping  $f$  from the set  $C$  of candidates onto the set  $P$  of positions. Since we have assumed that  $M \geq N$ , the mapping  $f$  assigns every candidate to a residency position. Let  $f_i$  denote the assigned specialty of candidate  $c_i$ , where  $f_i = 1$  if the candidate is assigned to family medicine, and  $f_i = 0$  otherwise. In the literature on matching mechanisms, including the well-studied house allocation problem (Abdulkadiroglu and Sonmez 1998), our candidates are called “agents,” our positions are called “houses,” and what we describe as an assignment is called an “allocation.” The house allocation literature generally assumes that the number of houses equals the number of agents. Here, we allow for the possibility that the number of positions differs from the number of candidates.

We further assume that each candidate  $c_i$  attains cardinal utility  $u_{Fi}$  if assigned to family medicine and cardinal utility  $u_{Hi}$  if assigned to the hospital-based specialty. Given his assignment  $f_i$ , the utility attained by candidate  $c_i$  can therefore be written as

$$u_i = f_i u_{Fi} + (1 - f_i) u_{Hi}.$$

While the matching literature generally specifies ordinal preference schemes, our use of cardinal utilities permits us to construct a social welfare function that directly addresses the tradeoff between equity and efficiency. (For an exception, see (Kelso and Crawford 1982).)

Conforming to the institutional setup in Spain, we restrict our analysis to one-sided matching systems. In contrast to two-sided matching mechanisms (Abdulkadiroglu, Pathak et al. 2009), the residency training positions in our framework are simply passive objects. They cannot express specific preferences for candidates, nor do their preferences enter into our normative analysis. One-sided mechanisms have also been studied in problems of school choice (Erdil and Ergin 2008).

We define the *private welfare function*  $U$  as

$$(1) \quad U = \sum_{i=1}^N w_i u_i$$

where the welfare weights  $w_i > 0$  attached to each candidate permit interpersonal comparisons and thus capture equity considerations. We will have more to say about the construction of these welfare weights below. Substituting our expression for each candidate's attained utility into (1), we get

$$(2) \quad U = \sum_{i=1}^N w_i (f_i u_{Fi} + (1 - f_i) u_{Hi}) = \sum_{i=1}^N f_i w_i \Delta u_i + K$$

where  $K = \sum_{i=1}^N w_i u_{Hi}$  is a constant term that does not depend on the assignment  $f$ . The

differential utility  $\Delta u_i = u_{Fi} - u_{Hi}$  is positive for a candidate who prefers family medicine and negative for one who prefers the hospital-based specialty.

Aside from its effects on private welfare, the assignment of candidates to training positions generates external social benefits. Specifically, we assume that there is an external social benefit to training a physician in family medicine, and that the magnitude of the externality varies among candidates. We define the *external benefit function*  $V$  as

$$(3) \quad V = \sum_{i=1}^N r_i f_i$$

where  $r_i$  measures the external benefit of moving candidate  $c_i$  from the hospital-based specialty to family medicine. We shall term  $r_i$  the *productivity* of candidate  $c_i$  to capture the idea that it is a characteristic intrinsic to each candidate but varies among candidates. The most productive candidates generate positive externalities. However, we allow for negative  $r_i$  to accommodate the possibility that transferring the least productive candidates into family medicine may generate negative externalities.

We can now frame our problem as maximizing the external benefit  $V$  subject to a constraint on private welfare  $U \geq U_0$ , where  $U_0$  is a constant. This is equivalent to maximize a *social welfare function* of the form

$$(4) \quad W = V + \lambda U$$

where  $\lambda > 0$  is a Lagrange multiplier. Substituting (2) and (3) into (4) gives the social welfare function

$$(5) \quad W = \sum_{i=1}^N f_i v_i + W_0$$

where  $v_i = r_i + \lambda w_i \Delta u_i$  and  $W_0 = \lambda K$ . The term  $v_i$  captures the net welfare effect of moving candidate  $c_i$  from the hospital-based specialty to family medicine.

### *First-Best Solution*

We can now characterize the first-best solution, that is, the assignment  $f: \mathbf{C} \rightarrow \mathbf{P}$  that maximizes the social welfare function  $W$  in (5).

Intuitively, it might seem that we could maximize social welfare simply by assigning those candidates with the highest values of  $v_i$  to the  $M_F$  positions available in family medicine. In fact, it is optimal to work in both directions. First, identify those candidates with *positive* values of  $v_i$ , assigning those with the highest values to family medicine and stopping if all  $M_F$  positions in that specialty are filled. Second, identify all those candidates with *negative* values of  $v_i$ , assigning those with the algebraically lowest values to the available positions in the hospital-based specialty, stopping when either there are no more candidates with negative values of  $v_i$  or when all  $M_H$  hospital-based positions have been filled. Third, assign those remaining candidates with the highest values of  $v_i$ , even if some values are negative, to all remaining positions in family medicine.

More formally, we partition the set  $\mathbf{C}$  of candidates into two mutually exclusive subsets based on the sign of each candidate's  $v_i$ . We denote  $\mathbf{C}_+ = \{c_i | v_i > 0\}$  and  $\mathbf{C}_- = \{c_i | v_i \leq 0\}$ , where

$C_+ \cup C_- = C$ . Let  $N_+ = |C_+|$  and  $N_- = |C_-|$  denote, respectively, the numbers of candidates in each subset, where  $N_+ + N_- = N$ . Then proceed in three steps.

*Step 1.1.* Assign those candidates  $c_i \in C_+$  with the highest values of  $v_i$  to family medicine. If this step fills all positions in that specialty (that is, if  $N_+ \geq M_F$ ), then the optimum has been attained, and the assignment of the remaining candidates will not affect social welfare  $W$ . Otherwise (that is, if  $N_+ < M_F$ ), continue to Step 1.2.

*Step 1.2.* Assign the remaining candidates  $c_i \in C_-$  with the algebraically lowest values of  $v_i$  to the subset of hospital-based positions. If this step assigns all remaining candidates (that is,  $N_- \leq M_H$ ), then the optimum has been attained. Otherwise (that is, if  $N_- > M_H$ ), continue to Step 1.3.

*Step 1.3.* Assign the remaining candidates  $c_i \in C$  to fill all remaining positions in family medicine.

So long as  $N < M$ , this three-step procedure is superior to simply assigning candidates with the highest values of  $v_i$  to family medicine. Only when  $N = M$  (that is, the number of candidates equals the number of positions) are the two solutions are equivalent. A trivial example makes the point. Suppose that there are just  $N = 2$  candidates with  $v_1 = 1$  and  $v_2 = -1$ , and that there are  $M = 4$  positions with  $M_F = M_H = 2$ . If we simply assigned the highest values of  $v_i$  to the two available positions in family medicine, then the social welfare (absent the constant term

$W_0$ ) would be  $W = \sum_{i=1}^2 f_i v_i = 0$ . But if we followed the above procedure, we would first assign

candidate  $c_1$  to family medicine (Step 1). Then we would assign candidate  $c_2$  to the hospital-based specialty (Step 2). With all candidates assigned a specialty, Step 3 would unnecessary and the social welfare (again, absent the constant term) would be  $W = 1$ . Note that if instead we began with only  $M = 2$  positions with  $M_F = M_H = 1$ , the two procedures would indeed be equivalent.

Figure 5 provides a richer example for a hypothetical set  $C$  of  $N = 90$  candidates who are to be assigned to a set  $P$  of 100 training positions, of which half are in family medicine and half are in the hospital-based specialty, that is,  $M_F = M_H = 50$ . Each candidate corresponds to a point in the  $(v, \Delta u)$  plane, where the subscripts  $i$  have been dropped for clarity. We have constructed the example so that  $N_+ = 35$  candidates have strictly positive values of  $v_i$  (that is, in the upper right and lower right quadrants), while the remaining  $N_- = 55$  candidates with non-positive values (the upper left and lower left quadrants). The crude correlation coefficient between  $v_i$  and  $\Delta u_i$  is  $-0.09$ .

The red-filled points represent the 40 candidates assigned to family medicine as follows. In Step 1.1, all  $N_+ = 35$  candidates in  $C_+$  are first assigned to family medicine. Since this step does not fill all  $M_F = 50$  positions in that specialty, we continue with Step 1.2, assigning those points with the  $M_H = 50$  lowest values of  $v_i$  (the open points in Figure 5) to the hospital-based specialty. Since this step still leaves some candidates unassigned (that is,  $N_- > M_H$ ), we continue to step 1.3, assigning the remaining  $N_- - M_H = 5$  candidates in  $C_-$  to family medicine. The example of Figure 5 shows how a central planner could take advantage of the excess supply of training positions to allow the hospital-based specialty to crowd out family medicine.

A central planner can achieve the first-best solution only if he knows the values of  $v_i$  of each candidate. While the values of the components  $r_i$ ,  $w_i$  and  $\lambda$  may be ascertainable, knowledge of each candidate's differential utility  $\Delta u_i$  is more problematic. Consider a candidate  $c_i$  who would prefer to be assigned to the hospital-based specialty, that is,  $\Delta u_i < 0$ . If he has a sufficiently high productivity  $r_i$ , he may end up assigned to family medicine. While such a candidate would want to reveal that he preferred the hospital-based specialty, he would nonetheless have an incentive to reveal an exaggerated negative value of  $\Delta u_i$ . An analogous argument shows that a low-productivity candidate who prefers family medicine would have an incentive to reveal an exaggerated positive value of  $\Delta u_i$ . Accordingly, in the case where the

central planner lacks information about the differential utilities  $\Delta u_i$  of the candidates, achieving the first-best solution is not *strategy proof* (Satterthwaite and Sonnenschein 1981). The breakdown of the strategy-proof characteristic in the case of cardinal utilities has been noted previously (Pathak 2011).

### *Priority Mechanisms and Second-Best Solutions*

In the MIR system, each candidate is ranked according to his combined performance in the nationwide exam and his medical school grade point average. The top-ranked candidate elects a residency training position among all available positions, after which the second highest ranked candidate elects a position among all the remaining positions, and the process continues sequentially. The MIR system is thus an example of a *priority mechanism*, also called “serial dictatorship” in the matching literature.

Any priority mechanism, including the MIR system, will result in an assignment  $f : C \rightarrow P$ . The question here is whether there is a priority mechanism that achieves the first-best optimum and, if not, what is the second-best optimum among all possible priority mechanisms.

Formally, assume that the priority mechanism orders the set of candidates  $C = \{c_i | i = 1, \dots, N\}$  so that candidate  $c_1$  chooses first,  $c_2$  chooses second, and so forth until  $c_N$  chooses last. Let  $P_i$  denote the subset of positions available to candidate  $c_i$  in the priority mechanism, where  $P_1 = P$ . If candidate  $c_i$  elects position  $f_i$ , then the choice set available to candidate  $c_{i+1}$  is  $P_{i+1} = P_i - \{f_i\}$ . If the choice set  $P_i$  consists solely of positions in only one of the two specialties, then candidate  $c_i$  and all remaining candidates  $\{c_{i+1}, \dots, c_N\}$  will have to choose that specialty. Otherwise, candidate  $c_i$  will choose  $f_i = 1$  if  $\Delta u_i \geq 0$ , and  $f_i = 0$  otherwise. To break ties, we've assumed that a candidate who is indifferent chooses family medicine.

It is a well-known result in the matching literature that a priority mechanism results in a Pareto-efficient allocation, provided that agents have strict preference orderings (Erdil and Ergin 2008). In our setup here, Pareto efficiency would mean that there is no other assignment  $f : C \rightarrow P$  that increases one candidate's utility without decreasing another's. Pareto efficiency

would likewise prevail for any priority mechanism, except in cases where at least one candidate is indifferent between specialties. To see why, suppose that there are only two candidates and two positions (that is,  $N = M = 2$ ) and that there is one available position in each specialty (that is,  $M_F = M_H = 1$ ). Assume further that  $\Delta u_1 = 0$  and  $\Delta u_2 > 0$ , that is, the first candidate is indifferent while the second candidate strictly prefers family medicine. When candidate  $c_1$  goes first, he breaks the tie and chooses family medicine, while candidate  $c_2$  will have to settle for the hospital-based specialty. Yet reversing the order of the two candidates would increase the utility of candidate  $c_2$  without decreasing the utility of candidate  $c_1$ .

However, as soon as we introduce the externality  $V$  or the social welfare weights  $w_i$ , we cannot guarantee that any priority mechanism achieves the first-best optimum of the social welfare function  $W$ . Just as we did for the first-best optimum, we characterize the second-best optimum among all priority mechanisms as a three-step procedure.

*Step 2.1.* Define  $C_F = \{c_i \mid \Delta u_i \geq 0\}$  as the subset of candidates who would prefer family medicine. Define  $C_{F+} = C_F \cap C_+$  as the subset of candidates who would prefer family and community medicine *and* have positive values of  $v_i$ . Rank those candidates in  $c_i \in C_{F+}$  in decreasing order of  $v_i$ , so that the candidate with the highest value of  $v_i$  is ranked first. If this step does not fill all  $M_F$  positions in family medicine, continue with step 2.2.

*Step 2.2.* Define  $C_H = C - C_F = \{c_i \mid \Delta u_i < 0\}$  as the subset of candidates who would prefer the hospital-based specialty. Place those candidates  $c_i \in C_H$  with the algebraically lowest values of  $v_i$  next in the ranking order, directly behind those already ranked in step 2.1. Let them continue to choose the hospital-based specialty, stopping when there are no more candidates (that is,  $|C_H| < M_H$ ) or no more hospital-based training positions (that is,  $|C_H| \geq M_H$ ).

*Step 2.3.* The order of the remaining candidates is arbitrary. If step 2.2 exhausts all candidates who would prefer the hospital-based specialty (that is,  $|C_H| < M_H$ ), then all remaining candidates will belong to  $C_F$  and elect family medicine. If step 2.2 exhausts all hospital-based

positions (that is,  $|C_H| \geq M_H$ ), then all remaining candidates will have to choose family medicine, regardless of their preferences.

Note that in the case where  $|C_H| \geq M_H$ , we have allowed the least socially valued candidates belonging to  $C_H$  to choose the hospital-based specialty until positions in that specialty are exhausted. This forces the more socially valued candidates in  $C_H$  to elect family medicine by default.

In a priority scheme, candidate  $c_i$  will choose the specialty with the higher utility so long as he has both types of training positions available to choose from. If training positions in only one of the two specialties remain open, he will have no alternative but to choose the remaining specialty. A priority scheme thus has two potential policy levers. It can identify candidates with high values of  $v_i$  who prefer family medicine and put them at the top of the ranking. It can also identify candidates with high values of  $v_i$  who would not choose family medicine and place them far enough in the back of the ranking so that, when their turn to choose comes up, there are no remaining slots in the other specialty to choose from.

Figure 6 exploits that same example as in Figure 5 to illustrate the second-best solution among all priority mechanisms. In Step 2.1, the 8 candidates in  $C_{F+} = C_F \cap C_+$  (the blue-filled points in the upper right quadrant) are placed at the top of the ranking and choose family medicine. In Step 2.2, all  $|C_H| = 60$  candidates who would prefer the hospital-based specialty are ranked next, starting the candidate in  $C_H$  with the algebraically lowest value of  $v_i$ , who corresponds to the points at the extreme lower left in the figure. This step fills up all  $M_H = 50$  positions in the hospital-based specialty, as represented by the open points in the figure (that is,  $|C_H| \geq M_H$ ). In Step 2.3, the remaining candidates (10 blue-filled points in the lower right quadrant and 22 blue-filled points in the upper left quadrant) elect the remaining  $90 - 8 - 50 = 32$  positions in family medicine. Thus, the 10 candidates with the highest social contribution are placed far enough back in the ranking that they are compelled to elect family medicine, contrary to their preferences.

In general, the second-best solution among priority schemes is not equivalent to the first-best solution. A sufficient condition is that the both the sets  $C_{F-} = C_F \cap C_-$  and  $C_{H+} = C_H \cap C_+$  be empty. In that case, all candidates with positive values of  $v_i$  would prefer family medicine, while all those with negative  $v_i$  would prefer the hospital-based specialty. In the  $(v, \Delta u)$  plane in Figures 5 and 6, there would be no points in the lower right or upper left quadrants. But this is clearly not a necessary condition. For example, we can relax the condition that  $C_{H+} = C_H \cap C_+$  (the lower right quadrant) is empty, making the additional assumption that  $M_H \leq |C_{H-}| = |C_H \cap C_-|$ , that is, there are enough candidates in the lower left quadrant to cover all the positions in the hospital-based specialty. In that case, we can first rank all candidates in  $C_{F+} = C_F \cap C_+$ , who will all choose family medicine. In the second step, we rank the candidates in  $C_{H-}$  in increasing order of  $v_i$  until the hospital-based positions are exhausted.

Achieving the second-best assignment among all priority rankings is subject to the same problems of revelation as the first-best assignment. If a central planner has to rely upon the candidates to reveal their differential utilities  $\Delta u_i$ , then those low-productivity candidates who prefer family medicine will have an incentive to exaggerate their positive values of  $\Delta u_i$ . Similarly, those high-productivity candidates who prefer the hospital-based specialty will have an incentive to exaggerate their negative values of  $\Delta u_i$ . Like the first best solution, the second-best assignment is not strategy proof.

### *The MIR Priority Mechanism*

The MIR selection system is a specific example of a priority mechanism in which the ranking of each candidate is based upon his combined score in the nationwide exam and his medical school grade point average. The candidate with the highest combined score goes first, while the candidate with the lowest score goes last.

Historically, the design of MIR system was motivated by three key considerations. First, the assignment mechanism was intentionally one-sided. This design element was imposed to avoid favoritism and political connections that might arise if the residency training programs had any say in the assignment of candidates. Second, the rank ordering of candidates was driven by a

strong sense of vertical equity. Those who perform better should be allowed to choose first. Third, the combined score on the national exam and medical school grade point average was regarded as a valid indicator of a candidate's productivity.

Within our conceptual framework, where the social value of candidate  $c_i$  is  $v_i = r_i + \lambda w_i \Delta u_i$ , the MIR's emphasis on vertical equity implies that a high-scoring candidate merits a high welfare weighting  $w_i$ . The MIR's design further implies that a high-scoring candidate has a high productivity. However, we see no convincing basis for the implicit assumption that a high-scoring candidate who prefers a hospital-based specialty such as neurosurgery would not be just as productive in family medicine. Thus, in our empirical evaluation of the MIR scheme below, we shall assume that a high-scoring candidate similarly has a high productivity  $r_i$  in family medicine.

But these are only secondary considerations. The fundamental problem with the MIR scheme is that the most productive candidates have assiduously avoided choosing family medicine. As noted in Figure 3, less than 2 percent of the top half of the ranking chose family medicine. While high-scoring, high-ranked candidates may indeed have high values of  $r_i$  and  $w_i$ , they clearly have negative values of  $\Delta u_i$ .

Figure 7 illustrates how a MIR priority mechanism would operate on the illustrative data in Figures 5 and 6. For simplicity, we have calculated each candidate's productivity as  $r_i = v_i - \Delta u_i$ , assuming that  $\lambda = w_i = 1$ , and then ranked candidates according their imputed values of  $r_i$ . The resulting correlation coefficient between  $r_i$  and  $\Delta u_i$  is  $-0.63$ .<sup>2</sup> The blue-filled points show the candidates assigned to family medicine, while the open points show those assigned to the hospital-based specialty. The five top-ranked candidates are labeled 1 through 5, while the five bottom-ranked candidates are labeled 86 through 90.

All candidates in the right lower quadrant (that is,  $C_{H+} = C_H \cap C_+$ ) are ranked high enough that they can exercise their preference for the hospital-based specialty until that specialty is filled. Moreover, all candidates in the upper quadrants (that is,  $C_F$ ) get to exercise their

---

<sup>2</sup> In our empirical analyses below, we will instead impute both  $r_i$  and  $w_i$  directly from the candidates' scores.

preference for family medicine. The remaining 10 candidates (the filled points in the left lower quadrant) with low values of  $v_i$  are forced to elect family medicine contrary to their preferences.

Thus, in contrast to the second-best priority mechanism, the MIR solution may assign the most productive candidates to the hospital-based specialty, while the least productive may end up in family medicine.

### *Supply Restriction*

Intuitively, it might seem that a selective reduction in the number of hospital-based positions would increase welfare, as more candidates would be shunted into family medicine. But this is not necessarily the case. In the implementation of the first-best solution or the second-best priority mechanism, the hospital-based specialty essentially serves as a reservoir for candidates with low values of  $v_i$ . Selective restriction of the hospital-based specialty may reduce the size of this reservoir, and thus reduce welfare. In the MIR priority mechanism, where the highest scoring, most productive candidates are also the highest ranked, the selective restriction of hospital-based training positions will not necessarily increase welfare either. Among those candidates who would otherwise prefer the hospital-based specialty, the restriction will force only the least productive into family medicine once the hospital-based specialty runs out of places.

Such supply restrictions operate only at the extensive margin. This point is illustrated in Figure 8, which shows the effect of imposing supply restrictions upon the MIR solution. The green filled points show the candidates assigned to family medicine under the MIR solution of Figure 7. The additional 10 yellow filled points show those candidates whose assignments are switched from the hospital-based specialty to family medicine when the number of hospital-based positions is reduced from  $M_H = 50$  to  $M_H = 40$ .

The supply restriction illustrated in Figure 8 resulted in a reduction in the total number of positions  $M$  from 100 to 90. As a result, the total number of positions equaled the total number  $N$  of candidates. More severe supply restrictions, which result in fewer positions than candidates (that is,  $N > M$ ), can also be incorporated into our conceptual framework. Appendix A provides the details.

### *Random Ordering*

In contrast to supply restrictions, random ordering of candidates will operate at the intensive margin. This is illustrated in Figure 9, which is based on the same illustrative data as in the previous Figures 5 through 8.

If we regard the MIR priority mechanism as our point of reference, a policy of random ordering can entail a considerable loss of private welfare. In our illustrative analyses of the MIR system in Figure 7 and the MIR system with supply restrictions in Figure 8, we set the welfare weights  $w_i = 1$  for simplicity. But in practice, the MIR system was motivated by the vertical equity notion that the highest performing candidates deserved to choose first. This would imply a strong correlation between the productivity  $r_i$  and the welfare weights  $w_i$ , which would amplify the private welfare losses resulting from random assignment.

The notion of vertical equity that underlies the MIR ordering based on a candidate's combined score is closely related to the concept of "justified envy" (Pathak 2011). Suppose that candidate  $c_1$  gets to choose a training position before candidate  $c_2$  does. We say that candidate  $c_1$  has "justified envy" of candidate  $c_2$  if the candidate  $c_1$  would derive higher utility from the other candidate's assignment  $f_2$  than his own assignment  $f_1$ . For all priority mechanisms, there are no such pairs of candidates that exhibit "justified envy." However, if we maintain the MIR ranking as our point of reference and then change the ranking to random ordering, then the new assignment does exhibit justified envy with respect to the original MIR ranking. In other words, if candidate  $c_1$  got to choose a training program before candidate  $c_2$  under the original MIR ranking and, as a result of a random ordering policy, candidate  $c_1$  would derive higher utility from the other candidate's assignment  $f_2$  than his own assignment  $f_1$ , then there is justified envy with respect to the original MIR ranking. Our use of welfare weights that reflect the original ranking capture this notion.

### *Policies to Change Utilities*

Consider a public policy variable  $X$  that influences candidates' preferences for different specialties but not their productivity. Without loss of generality, we assume that  $u_{Fi}(X)$  is strictly

increasing function of  $X$ , while  $u_{Hi}$  does not depend on  $X$ . As a result,  $\Delta u_i(X)$  is also an increasing function of  $X$ . We can therefore write the total social welfare function as

$$(6) \quad W(X) = \sum_{i=1}^N f_i(X) v_i(X) + W_0 - C(X)$$

where  $v_i(X) = r_i + \lambda w_i \Delta u_i(X)$  and  $C(X)$  is the cost of implementation as a function of  $X$ . In contrast to policies to restrict supply, changes in  $X$  operate at the intensive margin, affecting the differential utility values  $\Delta u_i$  of all candidates. We write  $f_i(X)$  because changing attributes will change the choices of some candidates under a priority mechanism.

Even if the differential utility  $\Delta u_i(X)$  is an increasing function of  $X$ , it is nonetheless possible that an increase in  $X$  could make some candidates worse off. Again, consider the MIR ranking with specific values  $X = X_0$  to be the reference case. Suppose that the top-ranked candidate  $c_1$  elects a hospital-based specialty (that is,  $\Delta u_1(X_0) < 0$ ). This choice leaves a position open in family medicine for a lower-ranked candidate  $c_i$ , who in fact prefers that specialty (that is,  $\Delta u_i(X_0) > 0$ ). With a sufficiently large increase in  $X$ , the top-ranked candidate  $c_1$  switches to family medicine, thus excluding the lower ranked candidate  $c_i$  from that specialty and forcing him instead to choose the hospital-based specialty. As a result of the increased attractiveness of family medicine, the higher ranked candidates crowd out the lower ranked candidates, who wanted residencies in family medicine in the first place.

#### 4. Data

Our principal database consisted of the individual assignments to residency training programs for all candidates participating in the 2011–2012 MIR nationwide competition (for short, the “2012 MIR registry”). This database was provided by Spain’s Ministry of Health, Social Services and Equality. For each candidate, the registry contained: the candidate’s national ranking (an ordinal number ranging from 1 up to the total number of participants); the residency program chosen (including medical specialty and training center); the candidate’s residential

postal code, sex, nationality, grade point average in medical school, and medical school attended, including foreign medical schools.

In the 2012 MIR competition, a total of 231 training centers offered residency positions in one or more of 44 specialties.<sup>3</sup> For hospital-based specialties, such as cardiology, neurology and urology, these training centers were located in 181 different hospitals. For non-hospital based specialties, such as family and community medicine, occupational medicine, public health, and some psychiatry training programs, we grouped the training centers according to 50 provinces throughout the country.

Together, these training centers offered a total of 2,527 distinct residency training programs, classified by specialty and center. With each residency training program offering multiple positions, there was a grand total of 6,555 available residency positions nationwide. Among these, 1,860 (28.4%) were residency positions in the specialty of family medicine. Initially, a total of 11,713 medical graduates passed the MIR exam in order to be eligible to opt for one of the 6,555 training positions. Of these candidates, 5,158 (44.0%) withdrew from the competition without choosing a residency, in many cases because their test score was so low that they had no chance of choosing their desired specialty or training center. That left 6,555 candidates for exactly as many training positions. A total of 1,716 (26.2%) of these participating candidates were foreign nationals.

We supplemented the 2012 MIR registry with additional data on the characteristics of the 231 training centers. These included information on hospitals' bed capacity and high-technology facilities, derived from the official National Catalog of Hospitals of Spain 2011. We also used Google Maps to create a 50×50 matrix of travel times between the capital of a candidate's province of residence and the capital of the province in which each training center was located.

We matched the 2012 MIR registry database with two cross-sectional surveys specifically constructed for this research project. The first was a survey of students in their final semester of medical school in Spain, administered in April 2011 (Harris, Gonzalez Lopez-Valcarcel et al. 2013). This survey provided us with the perceived values of seven key attributes of each

---

<sup>3</sup> While the Ministry of Health authorizes a total of 47 specialties, we excluded the school-based specialties of forensic medicine, sports medicine, and medical hydrology, which together accounted for only 149 candidates in the 2012 MIR competition. Rather than receiving salaries as medical residents, trainees in these three specialties pay tuitions to attend a professional school. They ranked at the bottom in the MIR competition.

specialty, as described in Appendix B, Table B1. These included: the probability of obtaining employment ( $X_1$ ), favorable working hours and working conditions ( $X_2$ ), recognition by patients ( $X_3$ ), prestige and recognition by colleagues ( $X_4$ ), possibilities for advancement and professional development ( $X_5$ ), average annual remuneration ( $X_6$ ), and percentage of income derived from private practice ( $X_7$ ). Among these attributes, we placed particular emphasis on professional prestige ( $X_4$ ) and outside income from private practice ( $X_7$ ). There is good evidence that professional prestige is an important driver of the decision not to pursue a career in primary care medicine, independent of income (Kolstad 2013). Moreover, in a nationalized healthcare system such as Spain, where employed physicians receive salaries negotiated through collective bargaining, the extent to which a physician can engage in outside private practice is a superior proxy for earnings (Harris, Gonzalez Lopez-Valcarcel et al. 2013). As a measure of earnings, outside income is more orthogonal to the other specialty characteristics than total earnings.

The second cross-sectional survey (for short, the “2012 post-MIR survey”) was performed in May 2012 on those candidates who had just been assigned to residency positions in the 2012 MIR competition. These candidates belonged to the same cohort that answered the first survey as medical students in 2011. We asked them not only about their actual specialty/center selection, but also about their preferred choice if they had been ranked first in the competition. We refer to the latter as a candidate’s *counterfactual* choice. This survey also included questions on gender, medical school, distance from home to the assigned residency training center, previous participation in prior MIR competitions, as well as respondents’ motivations for their current specialty/center selection. With our assurance of anonymity, respondents to this survey voluntarily provided their e-mail addresses, which permitted us to merge their responses with the principal database. We thus had additional data from this survey on 3,432 (or 52.4%) of the 6,555 candidates in the 2012 MIR registry. In our econometric models, to be described in Section 5 below, we dropped observations with missing values, leaving an estimation sample of 6,254 MIR candidates, of whom 3,189 could be matched with the survey database. Of these, a total of 3,117 had also answered the counterfactual question on their preferred specialty and training center.

Table 1 displays comparative statistics for the 2012 MIR registry and the 2012 post-MIR survey, including tests for comparison of group means. While there were no significant differences according to Spanish nationality or gender, those candidates responding to the post-

MIR survey performed better on the national exam and tended to choose positions closer to their residence.

For the 3,117 candidates with complete data in both databases, Table 2 compares their actual specialty choices in the MIR competition with their counterfactual ones. For a highly preferred specialty with a limited number of training positions, such as plastic surgery, only a small fraction of candidates were assigned to their preferred choices. By contrast, virtually every one of the candidates who preferred family medicine was assigned to his top choice. Those candidates who preferred cardiology, dermatology and plastic surgery tended to have high MIR rankings, so that few ended up assigned to family medicine.

In our analysis of the 2012 MIR registry, the dependent variable was the training program actually chosen by each candidate. In the 2012 post-MIR survey, the dependent variable was the training program that the candidate would have made if he had been top ranked (the counterfactual choice). In both analyses, the explanatory variables included the seven key attributes of the specialty; the characteristics of the training center (number of beds, and availability of positron emission tomography (PET) in the affiliated hospital); the characteristics of the candidate (gender, MIR ranking and nationality) interacted with the seven specialty attributes; the distance from the candidate's residence to the training center, measured in minutes of travel time between provincial capitals; and an indicator variable equal to 1 if the candidate's residence and the training center were located in the same province. The complete list of explanatory variables is shown in Appendix Table C1.

## **5. Discrete Choice Modeling**

Our estimation strategy was based upon a conditional logit model with random parameters, also called the "mixed logit model" (McFadden and Train 2000, Train 2009). To analyze the data from the 2012 MIR registry, we modified the model specifically to take into account the endogeneity of the choice set available to each candidate as his turn came up in the MIR sequence. We used the results of this model to simulate the policy options described in Section 6.

We also applied the mixed logit model to our 2012 post-MIR cross-sectional survey. In that case, respondents were asked to designate their preferred specialty under the counterfactual assumption that they had been ranked first, and therefore all candidates had the same, complete

choice set. We used the results of this model to test the validity of the axiom of Independence of Irrelevant Alternatives (IIA), as described below.

Generalizing the notation of Section 3 above, we let  $\mathbf{P} = \{p_j | j = 1, \dots, M\}$  denote the set of distinct residency training programs, each identified by a particular specialty and location, and let  $m_j$  denote the total number of training positions available in program  $p_j$ . Let  $\mathbf{C} = \{c_i | i = 1, \dots, N\}$  denote the set of candidates participating in the MIR sequential assignment process.

We first consider estimation of the mixed logit model with an endogenous choice set. We assume that the candidates are already ordered so that candidate  $c_1$  is first to elect a training program, while candidate  $c_N$  is last. Let  $\mathbf{P}_i$  denote the set of distinct residency training programs available to candidate  $c_i$  when it is his turn to elect a program, and let  $m_{ij}$  denote the number of remaining unassigned positions in training program  $p_j$  available to candidate  $c_i$  when his turn comes up. Since the top-ranked candidate  $c_1$  can choose any program, we have  $\mathbf{P}_1 = \mathbf{P}$  and  $m_{1j} = m_j$ . If candidate  $c_i$  chooses a position in training program  $p_j$ , the number of positions available in that training program is decremented by 1, that is,  $m_{i+1,j} = m_{ij} - 1$ , while the number of positions available in all other training programs remains unchanged, that is,  $m_{i+1,k} = m_{ik}$  for all  $k \neq j$ . If candidate  $c_i$  chooses the last available slot in training program  $p_j$ , that is,  $m_{ij} = 1$ , then  $\mathbf{P}_{i+1} = \mathbf{P}_i - \{p_j\}$  and that training program is dropped from the choice sets  $\mathbf{P}_{i+1}, \mathbf{P}_{i+2}, \dots$  of all remaining candidates.

Let  $y_i$  denote the training program chosen by candidate  $c_i$ . Given the unobserved parameter vector  $\beta$ , the probability that candidate  $c_i$  chooses training program  $y_i = p_j \in \mathbf{P}_i$  is given by the logit model

$$(7) \quad \Pr\{y_i = p_j \mid X_{ij}, \beta\} = \frac{e^{X_{ij}\beta}}{\sum_{p_k \in P_i} e^{X_{ik}\beta}} = L_{ij}(\beta)$$

where  $X_{ij}$  is a vector of observed characteristics that candidate  $c_i$  associates with program  $p_j$ .

Some elements of  $X_{ij}$  may depend only on the specific training program and not on the candidate, such as the particular specialty or the facilities of the training center. Other elements may depend on both the program and the candidate, such as the distance of the candidate's home province from the training center. Included in the latter category are interactions between a candidate's characteristics (e.g., gender, nationality) and a program's characteristics.

We further assume that the vector of unobserved parameters  $\beta$  has a multivariate normal distribution  $\beta \mid \mu, \Sigma \sim N(\mu, \Sigma)$  with density function  $\phi(\beta \mid \mu, \Sigma)$ . To simplify the notation, we let  $\theta = (\mu, \Sigma)$ . Conditional on  $\theta$ , the probability that candidate  $c_i$  choses training program  $y_i = p_j \in P_i$  is therefore given by<sup>4</sup>

$$(8) \quad \Pr\{y_i = p_j \mid X_{ij}, \theta\} = \int L_{ij}(\beta) \phi(\beta \mid \theta) d\beta = L_{ij}(\theta)$$

As noted by (Train 2009) and others, the parameters  $\theta = (\mu, \Sigma)$  can be estimated by maximum likelihood, where the integral in (9) can be computed by simulation. Let  $\hat{\theta}$  denote the maximum likelihood estimate of  $\theta$ , and consider a specific candidate  $c_i$ . Given his selected program  $y_i$ , the characteristics  $X_{ij}$ , and the estimate  $\hat{\theta}$ , the posterior density of  $\beta$  is given by

$$(9) \quad g(\beta \mid y_i = p_j, X_{ij}, \hat{\theta}) = \frac{L_{ij}(\beta) \phi(\beta \mid \hat{\theta})}{L_{ij}(\hat{\theta})}$$

---

<sup>4</sup> In (7) and (8), we have intentionally overloaded the notation  $L_{ij}(\cdot)$ , where  $L_{ij}(\beta)$  refers to the logit probability conditional on  $\beta$ , while  $L_{ij}(\theta)$  refers to the probability conditional on  $\theta$ .

We employed the *mixlogit* routine in Stata to estimate  $\hat{\theta}$  via maximum likelihood (StataCorp 2013). We imposed the additional restriction that covariance matrix  $\Sigma$  take the diagonal form  $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_K^2, 0, \dots, 0)$ , where the only first  $K$  elements of  $\beta$  were assumed to be random. In our subsequent simulations, to be described in the next section, we used the *mixlbeta* post-estimation routine (Hole 2007) to compute for each candidate  $c_i$  the posterior mean value of  $\beta$  in (10), which we denote by  $\hat{\beta}_i$ .

To analyze the 2012 post-MIR survey data, we assumed that each respondent made his counterfactual choice of preferred residency training program from the entire choice set  $\mathbf{P} = \{p_j | j = 1, \dots, M\}$ . Given the unobserved parameter  $\beta$ , the probability that candidate  $c_i$  preferred training program  $y_i = p_j \in \mathbf{P}$  is given by the logit model

$$(10) \quad \Pr\{y_i = p_j | X_{ij}, \beta\} = \frac{e^{X_{ij}\beta}}{\sum_{p_k \in \mathbf{P}} e^{X_{ik}\beta}} = L_{ij}(\beta)$$

Given  $\theta = (\mu, \Sigma)$ , the probability that candidate  $c_i$  preferred training program  $y_i = p_j \in \mathbf{P}$  is given by the same mixture of conditional probabilities shown in (8).

Application of the mixed logit model to the two different databases – the 2012 MIR registry and the 2012 post-MIR survey – permitted us to test the axiom of Independence of Irrelevant Alternatives (IIA) that is implicit in the logit model (8). In particular, for the 3,117 candidates who were represented in both databases, we compared the estimates  $\hat{\theta}$  derived from the 2012 MIR registry with those derived from the 2012 post-MIR survey. In the former case, the estimates are based on restricted choice sets  $\mathbf{P}_i$  dictated by the candidates' MIR ranking, while in the latter case, the estimates are based on the complete choice set  $\mathbf{P}$  of all training programs.

## 6. Simulation Strategies

In order to compare the impact of alternative public policies on private and social welfare, we used a simulation strategy based on the mixed logit model estimated from the entire

2012 MIR database. In each of the following scenarios, we used the estimated model parameters  $\hat{\beta}_i$  for each candidate  $c_i$  to predict his choice of training program from set of available programs  $P_i$ , proceeding sequentially from the first-ranked candidate  $c_1$  to the last candidate  $c_N$ . In effect, each simulation constitutes a realization of the MIR sequential selection process under different conditions, as determined by alternative public policies.

*Scenario 0 (Baseline)*: We simulated the MIR sequence based upon the original 2012 ranking, as determined by candidates' combined scores on the nationwide exam and medical school grade point average. We left the supply of training positions and candidates' preferences unchanged. This baseline case served as our reference for comparison.

*Scenario 1 (Random Assignment)*: We simulated the MIR sequence based upon the random ranking of candidates, rather than the actual 2012 ranking, leaving the supply of training positions and candidates' preferences unchanged. Specifically, we drew a pseudo-random number from a uniform distribution for each candidate, and then assigned to each candidate his resulting rank order statistic. We then ran 100 simulations of the MIR selection process, each time drawing a new random ordering of candidates.

*Scenario 2A (Concentrated Reduction in Supply)*: We maintained the original MIR ranking as well as the candidates' preferences, but reduced the total number of training positions by 20%, eliminating those training programs  $p_j$  with the fewest number of training positions  $m_j$ . This case was entirely deterministic with no random component, and was thus simulated once.

*Scenario 2B (Generalized Reduction in Supply)*: We maintained the original MIR ranking and candidates' preferences, but reduced the total number of training positions by 20%, eliminating positions at random from all training centers. To that end, we assigned to each training position a pseudo-random number from a uniform distribution, and then eliminated those training positions in the lowest 20% of the distribution. We then ran 100 simulations of the MIR selection process, each time randomly eliminating a different subset of training positions.

*Scenario 3 (Upgrading Family and Community Medicine)*: We maintained both the original MIR rank ordering and the supply of training positions, but varied the prestige and remuneration of family medicine as a specialty. Specifically, we repeatedly simulated the MIR

selection process for an array of different values of  $X_4$  (professional prestige) ranging from 0 to 10 in increments of 0.5, and of  $X_7$  (percentage of compensation derived from private practice) ranging in value from 0 to 100 in increments of 5. As noted in Appendix Table B1, the baseline valuations of these attributes for family and community medicine were  $X_4 = 3.92$  on a 10-point scale and  $X_7 = 13.2$  on a percentage scale. For all remaining specialties combined, the corresponding baseline valuations were  $X_4 = 6.30$  and  $X_7 = 39.3$ .

We refer to scenario 0 as the *baseline scenario*, while the remaining four scenarios as *alternative policy scenarios*.

For each of the above scenarios, including the baseline scenario, we computed an external benefit function,  $V = \frac{1}{N} \sum_{i=1}^N r_i f_i$ , normalized by the number  $N$  of candidates, where  $f_i = 1$  if and only if candidate  $c_i$  is assigned to family medicine, and where the candidate's productivity  $r_i$  was computed as a linear function of a candidate's MIR ranking. Specifically, let the candidate  $c_i$  with the highest MIR score have rank  $R_i = 1$ , while the candidate with the lowest score has rank  $R_i = N$ . We specified  $r_i = R_N - R_i + 1$ , so that the top-ranked candidate has a productivity of  $r_1 = R_N$  while the lowest ranked candidate has a productivity of  $r_N = 1$ .

For each scenario, including the baseline scenario, we also computed two private welfare functions,  $U = \frac{1}{N} \sum_{i=1}^N r_i u_i$  and  $\bar{U} = \frac{1}{N} \sum_{i=1}^N \bar{r} u_i$ , where  $u_i = \max_j \{u_{ij}\}$  represents the maximum utility attained by candidate  $c_i$  among all training programs  $p_j$  available in his choice set. In our formulation of  $U$ , we adhered to the notion of vertical equity inherent in the MIR scheme and set the welfare weights equal to the candidate's productivity, that is,  $w_i = r_i$  for all  $i = 1, \dots, N$ . By contrast, in our formulation of  $\bar{U}$ , we assumed equal welfare weights  $w_i = \bar{r}$ , where  $\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i$  is the mean productivity. We shall call the former *productivity-driven weights* and the latter *homogenous weights*.

For each alternative policy scenario  $s = 1, 2A, 2B, 3$ , we then computed the differences  $\Delta V_{(s)} = V_{(s)} - V_{(0)}$ ,  $\Delta U_{(s)} = U_{(s)} - U_{(0)}$ , and  $\Delta \bar{U}_{(s)} = \bar{U}_{(s)} - \bar{U}_{(0)}$ , where the subscripts in parentheses refer to scenarios and where the subscript (0) refers to the baseline scenario.

Finally, we displayed the results of simulations as scatterplots, in which each individual candidate  $c_i$  corresponded to a single point in the plane. For each scenario  $s = 0, 1, 2A, 2B, 3$ , including the baseline scenario, we graphed the candidate's maximum attained utility  $u_{i(s)}$  versus his original MIR ranking  $R_i$ . Then, for each alternative policy scenario  $s = 1, 2A, 2B, 3$ , we graphed the change in the candidate's utility  $\Delta u_{i(s)} = u_{i(s)} - u_{i(0)}$  from the baseline versus his original MIR ranking  $R_i$ .

## 7. Results: Discrete Choice Models

We estimated discrete choice three models, which we designate as I, II and III, respectively. The complete results are reported in Appendix Table C1. Model I covered the 6,254 candidates in the 2012 MIR registry who had complete data for all variables. The dependent variable was the training program actually chosen by each candidate in the 2012 MIR competition, while each candidate's choice set consisted of those training programs still available when his turn came up in the MIR sequence. Model II was a replication of Model I restricted to the subsample of 3,117 candidates who fully responded to the 2012 post-MIR survey. Model III covered the same subsample of 3,117 respondents to the 2012 post-MIR survey, but with the counterfactual dependent variable – namely, the training program each candidate would have chosen if he had been ranked first. Each candidate's choice set consisted of all available training programs.

In each of the three models, we permitted only a subset of the explanatory variables to have random coefficients. The estimated mean values of the coefficients of all explanatory variables are listed in the upper portion of Table C1 under the heading “Mean ( $\mu$ ),” while the

estimated standard deviations for those explanatory variables with random coefficients are listed in the lower portion under the heading “Standard Deviation ( $\sigma$ ).”<sup>5</sup>

#### *Model I: 2012 MIR Registry*

Model I correctly predicted 34 percent of the specialty choices of the 6,254 candidates. However, with respect to the choice of family and community medicine, the concordance between observed and predicted was 90 percent. In Model I, all location-related variables and hospital characteristics were highly significant and had the expected signs. Distance thus imposes a cost that candidates took into account in their decisions, while remaining in one’s province of residence conferred an additional positive component to a candidate’s utility.

In Model I, all specialty attributes had significant coefficients except for annual remuneration ( $X_6$ ). When normalized by their sample standard deviations, three attributes had the largest quantitative effects: the probability of obtaining employment ( $X_1$ ), prestige among colleagues ( $X_7$ ), and the percent of compensation earned in private practice ( $X_7$ ). Moreover, three attributes displayed significant variability among candidates, as indicated by their estimated standard deviations: the probability of obtaining employment ( $X_1$ ), the recognition by patients ( $X_3$ ) and the percent of compensation earned in private practice ( $X_7$ ).

With respect to interactions between specialty attributes and the characteristics of the candidates, Spanish nationals valued the probability of obtaining employment ( $X_1$ ) more than foreign candidates. Higher-ranked candidates (with lower absolute values of  $R_i$ ) valued this attribute more than lower-ranked candidates. Women, foreign candidates and lower-ranked candidates placed more value on prestige among colleagues ( $X_4$ ) than their counterparts. Women, Spanish nationals and higher-ranked candidates attached more value to income from private practice ( $X_7$ ) than their respective counterparts.

#### *Models II and III: The 2012 Post-MIR Survey*

In Appendix Table C1, none of the estimated coefficients of Model II differed significantly from those of Model I, a finding that points to the absence of selection bias among

---

<sup>5</sup> The decision as to which explanatory variables had random coefficients was empirically based. We first estimated a more general model in which all explanatory variables had random coefficients, and then restricted the random coefficients of the explanatory variables to those with significant standard deviations at the 5% level.

those who responded to the 2012 post-MIR survey.<sup>6</sup> Both Models II and III correctly predicted 31 percent of the specialty choices of the 3,711 candidates. With respect to the choice of family and community medicine, the concordance between observed and predicted was 87 and 98 percent for Models II and III, respectively.

In Model III, each candidate's choice set corresponded to all available training positions nationwide. In Model II, by contrast, each candidate's choice set was restricted to those training positions available when his turn came up in the MIR competition. Comparison of the parameter estimates of the two models thus constitutes a test of the axiom of the Independence of Irrelevant Alternatives.

In Appendix Table C1, 25 out of the 33 estimated coefficients showed no significant difference between Models II and III. Among the 8 coefficients with significant differences, two notable instances were the number of positions available in the training program, and the interaction between Spanish nationality and the probability of obtaining employment ( $X_1$ ). While the size of the training program had a positive effect in both Models II and III, it mattered even more when a candidate's choices were restricted by his rank ordering in the MIR scheme (Model II). For Spanish medical graduates, job security likewise mattered significantly more when a candidate's choices were restricted (Model II).

In Table 3, we compare the coefficient estimates derived from Models II and III for all interactions between the seven medical specialty attributes ( $X_1$  through  $X_7$ ) and the candidate's MIR ranking. In Model III, where there were no restrictions on choice, a candidate's ranking in the MIR scheme functioned solely as a proxy for his skill and productivity. In our econometric model, MIR ranking was defined so that lower ranked candidates with inferior national exam scores and poor medical school grades had higher values of  $R_i$ , while the top-ranked candidate had a MIR ranking of  $R_1 = 1$ . Accordingly, the primarily negative signs of the coefficients in Model III imply that the more qualified candidates actually attached greater value than their less productive counterparts to all specialty characteristics except recognition by patients ( $X_4$ ).

---

<sup>6</sup> Comparison of coefficients was based on the Z-statistic:  $Z = (b_i - b_{ii}) / \sqrt{s_i^2 + s_{ii}^2}$ , where  $b_i$  and  $b_{ii}$  are the estimated parameters from the two models, while  $s_i$  and  $s_{ii}$  are their respective standard errors.

In Model II, where a candidate's MIR ranking served not only as a proxy for a candidate's qualifications but also as a determinant of his choice set, the coefficients of the interaction terms were generally reduced in absolute value or even lost statistical significance. In the case of prestige among colleagues ( $X_7$ ), the coefficient was actually reversed. Less qualified candidates, once they were compelled to choose among the slim pickings available to those at the bottom of the MIR rankings, valued prestige among their colleagues even *more* than their top ranked counterparts.

## 8. Results: Simulations

Table 4 summarizes the main results of our simulations, including the criterion functions  $\Delta V$ ,  $\Delta \bar{U}$ , and  $\Delta U$ . In addition, for each scenario  $s = 0, 1, 2A, 2B, 3$ , including the baseline scenario, Figures D1–D5 in Appendix D plot the maximum utility  $u_{i(s)}$  attained by each of the 6,255 MIR candidates in relation his original MIR ranking  $R_i$ . For each alternative policy scenario  $s = 1, 2A, 2B, 3$ , Figures 10–13 plot the change in utility  $\Delta u_{i(s)} = u_{i(s)} - u_{i(0)}$  from the baseline in relation to the candidate's original MIR ranking  $R_i$ . For scenarios 1 and 2B, which required multiple replications, we plot the results of one of the draws as an illustration.

For each of the alternative policy scenarios  $s = 1, 2A, 2B, 3$ , Table 4 shows an improvement in the external benefit  $V_{(s)}$  compared to its baseline value  $V_{(0)}$ . The improvement was largest for policy scenario 3 (Upgrading Family and Community Medicine).

In Table 4, random assignment of candidates (scenario 1) improved the welfare criterion  $\bar{U}$  so long as we used homogeneous weights. With the use of productivity-driven weights, however, the criterion function  $U$  diminished markedly. In Figure 10, the bottom ranked candidates with high values of  $R_i$  experienced utility gains, while the top ranked candidates with low values of  $R_i$  experienced utility losses. As indicated by the blue points in the figure, a large proportion of these top ranked candidates were compelled to elect family medicine when they were randomly placed at the back of the queue. Accordingly, when interpersonal comparisons are based on the notion that the most talented medical students deserve to make their specialty

choices first, the random assignment policy results in a substantial offsetting equity loss. The narrow credibility intervals of the criterion functions for this policy indicate that variations in the specific random ordering of candidates would not alter this conclusion.

Policies intended to reduce the supply of training positions (scenarios 2A and 2B) resulted in modest improvements in the external benefit indicator  $V$ , but substantial welfare losses for both the homogeneous ( $U$ ) and productivity-driven ( $\bar{U}$ ) criterion functions. As shown in Figures 11 and 12, all candidates incurred losses in utility as a result of restrictions in their available choices, while those candidates who were excluded from any training position incurred a substantial utility losses. The improvements in  $V$  arose from the fact that some candidates at the margin were compelled to choose family medicine. In Table 4, the narrow credibility intervals in the row corresponding to alternative policy 2B indicate that variations in the specific training positions to be excluded would not alter these conclusions.

Upgrading family and community medicine (scenario 3) was the only policy with the potential for improvements in both equity and efficiency. Table 4 shows the effects of a policy that would increase both  $X_4$  (prestige among colleagues) and  $X_7$  (proportion of income earned in private practice) from the current mean values for family medicine to the mean values of all other specialties. Figure 13 shows improvements in utility among top ranked candidates who have now elected family and community medicine for its superior attributes. In the figure, 11 percent of the candidates experienced a net *loss* in utility compared to the baseline scenario, that is,  $\Delta u_{i(3)} = u_{i(3)} - u_{i(0)} < 0$ . Nearly all of these candidates had chosen family medicine in the baseline scenario. They were deprived of their preferred specialty choice – or, in a few cases, their preferred training location – once family medicine became more attractive to the top ranked candidates.

Figures 14 and 15 plot the simulation results for different combinations of  $X_4$  and  $X_7$ . The yellow-filled point in each figure corresponds to  $X_4 = 6.30$  on a 10-point scale and  $X_7 = 39.3$  on a percentage scale, which were assumed in Table 4, while the intersection of the solid horizontal and vertical lines corresponds to the status quo, where  $\Delta V_{(3)} = \Delta U_{(3)} = 0$ . The figures illustrate the tradeoffs between the two attributes in attracting candidates to family and community

medicine. For example, to achieve approximately the same external benefit criterion  $\Delta V_{(3)} = 859$  and the same productivity-driven welfare criterion  $\Delta U_{(3)} = 3835$  as in Table 4, we could have selected the alternative values  $X_4 = 7.4$  and  $X_7 = 25$  percent. This gives a local marginal rate of transformation equal to  $-(25 - 39.3)/(7.4 - 6.3) = 13$  percentage points of income from private practice for every 1 point on a 10-point scale of prestige among colleagues.

## 9. Discussion and Conclusions

We modeled Spain's MIR system of assigning residency training positions as a one-sided matching mechanism based on priority rankings, also called "serial dictatorship." Within this theoretical framework, we formulated a social welfare criterion that explicitly took account of the tradeoff between two key components: the external social benefit of having highly productive candidates elect a residency in family and community medicine; and the vertical equity gains from permitting the most productive candidates to have the highest priority in choosing a training program. We compared the MIR system to the optimal allocation of residency positions, as well as the second-best among all sequential allocation mechanisms. We theoretically characterized alternative policies for improving the MIR allocation, including random ranking of candidates, restrictions on the supply of residency positions, and improvements in the perceived attributes of family and community medicine.

We then applied our theoretical framework to real data. We formulated and estimated a multinomial logit model with random coefficients to capture the preferences of candidates participating in the 2012 MIR competition. We then used the estimated parameters of this model to simulate the effects of alternative public policies. We found that random ranking of candidates (policy scenario 1) resulted in a relatively small increase in the external social benefit of training highly productive candidates in family medicine, but at the same time a substantial equity loss. Moreover, policies aimed at reducing the number of training positions (scenarios 2A and 2B) improved the external social benefit at the extensive margin, but likewise resulted in a substantial loss of welfare, as candidates at the bottom of the MIR ranking were locked out of any training position at all. By contrast, policies that increased two key attributes of family medicine –

professional prestige and compensation from private practice – increased the external social benefit at the intensive margin without any countervailing loss of welfare.

The four scenarios that we considered do not necessarily cover the full range of available policy options, but they do reflect realistic choices given current political constraints in Spain. We selected policies scenarios 2A and 2B, in particular, because the Spanish government has already made a decision to reduce the total number of residency training positions as part of cutbacks in response to the financial crisis. We chose policy scenario 3 to capture a range of recent proposals to fundamentally change the nature of work of primary care physicians in Spain (Gonzalez Lopez-Valcarcel, Barber Perez et al. 2011). In modeling fundamental changes in the perception of family and community medicine, we focused on two key attributes: prestige among colleagues ( $X_4$ ), and the proportion of income derived from private practice ( $X_7$ ). In addition to job security ( $X_1$ ), these factors had been identified as critical in medical school students' assessments of careers in alternative specialties (Harris, Gonzalez Lopez-Valcarcel et al. 2013).

We chose policy scenario 1 (random ranking of candidates) not because it is currently politically feasible, but because it serves as an important theoretical benchmark. Our empirical finding that random ranking resulted in a substantial equity loss serves as a cautionary note to the general perception that “random serial dictatorship” and its variants necessarily result in fair allocations (Budish, Che et al. 2013). Random assignment may be regarded as extremely unfair when there is a strong perception that rankings should be based on merit, performance or similar criteria.

Among the policy alternatives that we did not consider is a secondary market, where subsets of candidates could mutually trade residency positions after the completion of the formal assignment process. This possibility is related to Gale's “top trading cycles” algorithm (Shapley and Scarf 1974). Even more radical would be a pre-MIR market, in which candidates with high exam scores could sell their rankings to those at the bottom and then take the national exam the following year. This practice – which we might call “I sell my turn” – bears a resemblance to the decision to give up one's reserved seat on an airline flight in return for compensation and a later reservation.

For the subset of MIR candidates who responded to our 2012 post-MIR survey, we compared the results of a multinomial logit model with constrained choice sets (Model II) to the same model specification with unconstrained choice sets (Model III). This comparison allowed

us to test for violations of the axiom of Independence of Irrelevant Alternatives in a manner different from that described by Berry and colleagues, who had data on consumers' first and second choices of automobiles (Berry, Levinsohn et al. 2004). Our approach of restricting choice sets resembles that of Machado and colleagues, who used the revealed preferences of MIR candidates to make inferences about the relative quality of hospitals' training programs in Spain (Machado, Mora et al. 2012). It differs from that of Timmins and Murdock, who addressed the problem of congestion in travel cost models by modifying the specification of the utility function (Timmins and Murdock 2007).

We found, in particular, that the interaction between MIR ranking and prestige among colleagues ( $X_4$ ) reversed sign in the counterfactual scenario, while the interaction with the proportion of income from private practice ( $X_7$ ) did not (Table 3). Under the constraints of the MIR system, those candidates near the bottom of the ranking attached more importance to prestige, while in a counterfactual world without such constraints, those near the top attached more importance to this attribute. In the real world where the highest scorers go first, professional prestige becomes a marker of exclusivity that is desired by those with fewer options. In a counterfactual world without any restrictions on choice, professional prestige is desired solely by those with the best qualifications. In essence, restricting choice creates an externality in which those at the bottom, who have fewer choices, want what those at the top already have (Bikhchandani, Hirschleifer et al. 1992).

The possibility that a candidate's ranking in MIR priority mechanism itself influences his underlying preferences would complicate our analysis of public policies such as random ranking. In effect, our simulations are designed to capture local effects, in which policies do not change candidate's underlying preferences, as embodied in the parameters  $\beta$ , but only the explanatory variables ( $X, Z$ ) in our choice model (equation 7). It is conceivable that changes, for example, in the organization of family and community medicine may result in global changes in candidates' preferences.

Our study has a number of important limitations. First, we measured the two key components of our social welfare criterion, namely, the external benefit  $V$  and the private welfare  $U$ . But we had no obvious strategy for identifying the multiplier  $\lambda$  in the overall social welfare  $W = V + \lambda U$ . Rather than speculating about alternative values of  $\lambda$ , we chose simply to

display the two key components separately (Table 4). Nonetheless, we believe that the comparisons are so stark that efforts to identify  $\lambda$  through surveys of physicians and policymakers would add little.

In our calculations of the welfare effects of restrictions in supply (policy scenarios 2A and 2B), we arbitrarily assigned a utility  $u_i = 0$  to each candidate who was excluded from any residency training program. This assumption of “no life outside MIR” ignores any residual utility from serving as a non-physician. Nor does it account for the possibility that an excluded candidate could again apply for a residency position the following year. Again, we believe that our comparisons in Table 4 are so stark that efforts to identify the residual utility are unlikely to influence our conclusions.

In our simulations, we used the posterior mean values of the coefficients  $\hat{\beta}_i$ , based on the posterior density function given in equation (9) above. In a simulation study of school choice, Hastings and colleagues adopted a similar strategy (Hastings, Kane et al. 2010). Since we did not simulated repeated draws from the posterior density, the results reported in Table 5 understate the degree of uncertainty in our results.<sup>7</sup>

In our simulations, we used each candidate’s ranking as a proxy for his productivity  $r_i$ . It is arguable that scores on a national exam and performance in medical school are unreliable indicators of long-run productivity, and therefore don’t merit inclusion in our calculation of the external benefit  $V$  from assigning candidates to family and community medicine. The available evidence, however, suggests that one’s score on a national exam taken during medical school is a stable, reliable indicator of productivity over the long run (Epstein, Nicholson et al. 2013).

In our analysis of policies to upgrade family and community medicine, we focused sharply on prestige among colleagues ( $X_4$ ) and the proportion of income from private practice ( $X_7$ ). We ignored total remuneration ( $X_6$ ), which in fact had a borderline significant negative sign in Model I (Table C1,  $P = 0.075$ ). Spanish medical students appear to have limited quantitative

---

<sup>7</sup> We did not adopt the procedure suggested by Revelt and Train (2000), where the conditional density of  $\beta$  given the selected program  $y_i$ , the characteristics  $X_{ij}$ , and the estimate  $\hat{\theta}$  can be approximated by repeatedly drawing  $\beta^r$  from the multivariate normal distribution  $N(\hat{\mu}, \hat{\Sigma})$  and then weighting its value by  $L_{ij}(\beta^r) / \sum_{p_i \in P_i} L_{ik}(\beta^r)$ .

knowledge of the earnings of different specialties 10–15 years out of medical school, a phenomenon that has also been observed in the U.S. (Nicholson 2005). In a national system where nearly all physicians are employees, the proportion of earnings from private practice appears to a superior indicator of earnings that is more orthogonal to other specialty characteristics (Harris, Gonzalez Lopez-Valcarcel et al. 2013). Similar problems in identifying the orthogonal component of earnings have been encountered in studies of predominantly fee-for-service payment systems (Gagne and Leger 2005).

Our welfare analysis did not take into account the costs of implementation of alternative policies, as reflected in the term  $C(X)$  in equation (6) above. This drawback is especially important for scenario 3, where we studied policies aimed at enhancing the professional prestige and remuneration of practitioners of family and community medicine. Improving the professional prestige of family doctors will likely entail the creation of top-level teaching positions in medical schools, as well as the redirection of research and development funds toward the study of primary care. At present, only a small minority of family doctors have their own private practices or receive compensation for being on-call at private hospitals. New incentives may be necessary to stimulate the formation of physician-owned health centers that would eventually assume responsibility for the primary care of the general population under contract with the Spain's National Health Service. To the extent that income from private practice served as a proxy for total compensation, our results suggest that increased salaries of family doctors may be an effective means to attract the most productive young physicians to family medicine. In the long term, enhancement of professional prestige and increased opportunities for remuneration from private practice may require a reformulation of the nature of work of the family physician, as well as the role of community health centers not only in Spain, but also in other healthcare systems.

We conclude that policies designed to increase the prestige and remuneration of practitioners of family and community medicine have the potential to be more efficient and equitable than other alternatives. Other policies that we studied here resulted in moderate improvements in the number of productive practitioners of family medicine, but at substantial costs in terms of equity.

**Funding**

This project was funded by the Spanish Ministry of Science and Innovation through Grant ECO2010–21558 to BGLV as the principal investigator. The funder had no influence in the conduct of this study or the drafting of this manuscript.

**Acknowledgments**

We gratefully acknowledge the Ministry of Health, Social Services and Equality, General Sub-Directorate on Professional Planning, Spain, for providing us the 2012 MIR registry data. We acknowledge the many medical students and MIR candidates who responded to our 2011 and 2012 surveys. We thank Jaime Pinilla Domínguez for technical assistance. Finally, we thank the participants in workshops and seminars that we have given concerning this research, including: the Pontificia Universidad Católica, Chile (JEH), Universidad Pública de Navarra (BGLV), Universidad de Vigo (BGLV), and the Grupo Evaluación de Políticas y Servicios de Salud, Asociación de Economía de la Salud (EvaluAES), Valencia (VO). The opinions expressed in this paper are ours and ours alone.

**Conflicts of Interest**

We have none to report.

**Author Contributions**

All authors (JEH, BGLV, PB, VO) participated in the conceptualization and the design of the study, as well as the analysis and interpretation of the data. JEH and BGLV were principally responsible for writing the manuscript. All authors have reviewed and approved the final draft of this working paper.

## References

Abdulkadiroglu, A., P. A. Pathak and A. E. Roth (2009). "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." American Economic Review **99**(5): 1954-1978.

Abdulkadiroglu, A. and T. Sonmez (1998). "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems." Econometrica **66**(3): 689-701.

Abdulkadiroğlu, A. and T. Sönmez (1998). "Random serial dictatorship and the core from random endowments in house allocation problems." Econometrica **66**(3): 689-701.

Abdulkadiroğlu, A. and T. Sönmez (2003). "School choice: A mechanism design approach." American economic review: 729-747.

Barber, P. and B. Gonzalez Lopez-Valcarcel (2010). "Forecasting the need for medical specialists in Spain: application of a system dynamics model." Human Resources for Health **8**(24).

Barber, P. and B. G. Lopez-Valcarcel (2010). "Forecasting the need for medical specialists in Spain: application of a system dynamics model." Hum Resour Health **8**: 24.

Berry, S., J. Levinsohn and A. Pakes (2004). "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." Journal of Political Economy **112**(1): 68-105.

Bikhchandani, S., D. Hirschleifer and I. Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." Journal of Political Economy **100**(5): 992-1026.

Bodenheimer, T. (2006). "Primary care--will it survive?" N Engl J Med **355**(9): 861-864.

Budish, E., Y.-K. Che, F. Kojima and P. Milgrom (2013). "Designing Random Allocation Mechanisms: Theory and Applications." American Economic Review **103**(2): 585-623.

Cheng, T. C., A. Scott, S. H. Jeon, G. Kalb, J. Humphreys and C. Joyce (2011). "What Factors Influence the Earnings of General Practitioners and Medical Specialists? Evidence from the Medicine in Australia: Balancing Employment and Life Survey." Health Econ.

Colwill, J. M., J. M. Cultice and R. L. Kruse (2008). "Will generalist physician supply meet demands of an increasing and aging population?" Health Aff (Millwood) **27**(3): w232-241.

de Bakker, D. H., J. N. Struijs, C. B. Baan, J. Raams, J. E. de Wildt, H. J. Vrijhoef and F. T. Schut (2012). "Early results from adoption of bundled payment for diabetes care in the Netherlands show improvement in care coordination." Health Aff (Millwood) **31**(2): 426-433.

Dorsey, E. R., D. Jarjoura and G. W. Rutecki (2003). "Influence of controllable lifestyle on recent trends in specialty choice by US medical students." JAMA **290**(9): 1173-1178.

Epstein, A. j., S. Nicholson and D. A. Asch (2013). The Production of and Market for New Physicians' Skill, National Bureau of Economic Research, Working Paper No. 18678, January.

Erdil, A. and H. Ergin (2008). "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice." American Economic Review **98**(3): 669-689.

Ergin, H. I. (2002). "Efficient resource allocation on the basis of priorities." Econometrica **70**(6): 2489-2497.

Gagne, R. and P. T. Leger (2005). "Determinants of physicians' decisions to specialize." Health Econ **14**(7): 721-735.

Gonzalez Lopez-Valcarcel, B. and P. Barber Perez (2012). "[Health workforce planning and training, with emphasis on primary care. SESPAS Report 2012]." Gac Sanit **26 Suppl 1**: 46-51.

Gonzalez Lopez-Valcarcel, B., P. Barber Perez and V. Ortun Rubio (2011). ¿Por qué los médicos huyen y rehúyen la Medicina de Familia? Datos y claves sobre el problema en busca de soluciones. La refundación de la Atención Primaria. V. Ortún. Madrid, Springer Healthcare.

Gonzalez Lopez-Valcarcel, B., V. Ortun, P. Barber and J. E. Harris (2013). "[Important differences between faculties of medicine. Implications for family and community medicine.]" Aten Primaria.

Goroll, A. H., R. A. Berenson, S. C. Schoenbaum and L. B. Gardner (2007). "Fundamental reform of payment for adult primary care: comprehensive payment for comprehensive care." Journal of General Internal Medicine **22**(3): 410-415.

Harris, J. E., B. Gonzalez Lopez-Valcarcel, V. Ortun and P. Barber (2013). "Specialty choice in times of economic crisis: a cross-sectional survey of Spanish medical students." BMJ Open **3**(2).

Hastings, J. S., T. J. Kane and D. O. Staiger (2010). Heterogeneous Preferences and the Efficacy of Public School Choice, New Haven CT, Yale University Department of Economics, Working Paper.

Hole, A. (2007). "Fitting mixed logit models by using maximum simulated likelihood." The Stata Journal **7**(3): 388-401.

Huibers, L., P. Giesen, M. Wensing and R. Grol (2009). "Out-of-hours care in western countries: assessment of different organizational models." BMC Health Serv Res **9**: 105.

Kelso, A. S., Jr., and V. P. Crawford (1982). "Job matching, coalition formation, and gross substitutes." Econometrica **50**(6): 1483-1504.

Kolstad, J. T. (2013). Information and Quality when Motivation is Intrinsic: Evidence from Surgeon Report Cards, National Bureau of Economic Research, Working Paper No. 18804, February.

Krueger, K. J. and E. C. Halperin (2010). "Perspective: Paying physicians to be on call: a challenge for academic medicine." Acad Med **85**(12): 1840-1844.

Lopez-Valcarcel, B. G., V. Ortun, P. Barber, J. E. Harris and B. Garcia (2013). "Ranking Spain's Medical Schools by their performance in the national residency examination." Rev Clin Esp **213**(9): 428-434.

Machado, M. P., R. Mora and A. Romero-Medina (2012). "Can We Infer Hospital Quality from Medical Graduates' Residency Choices?" Journal of the European Economic Association **10**(6): 1400–1424.

Mariolis, A., C. Mihas, A. Alevizos, V. Gizlis, T. Mariolis, K. Marayiannis, Y. Tountas, C. Stefanadis, A. Philalithis and G. Creatsas (2007). "General Practice as a career choice among undergraduate medical students in Greece." BMC Med Educ **7**: 15.

McFadden, D. and K. Train (2000). "Mixed MNL Models for Discrete Response." Journal of Applied Econometrics **15**: 447-470.

Nicholson, S. (2005). "How Much Do Medical Students Know about Physician Income?" Journal of Human Resources **40**(1): 100-114.

Ortun, V., B. Gonzalez Lopez-Valcarcel and P. Barber Perez (2008). "[Determinants of physician compensation]." Med Clin (Barc) **131**(5): 180-183.

Pathak, P. A. (2011). "The Mechanism Design Approach to Student Assignment." Annual Reviews of Economics **3**(1): 513-536.

Rosenblatt, R. A., C. H. Andrilla, T. Curtin and L. G. Hart (2006). "Shortages of medical personnel at community health centers: implications for planned expansion." JAMA **295**(9): 1042-1049.

Roth, A. E. and E. Peranson (1999). "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." American Economic Review **89**(4): 748-780.

Satterthwaite, M. A. and H. Sonnenschein (1981). "Strategy-Proof Allocation Mechanisms at Differentiable Points." Review of Economic Studies **48**(4): 587-597.

Scott, J., A. Larson, F. Jefferies and B. Veenendaal (2006). "Small-area estimates of general practice workforce shortage in rural and remote Western Australia." Aust J Rural Health **14**(5): 209-213.

Servicio Público de Empleo Estatal (SEPE) (2013). Información mensual de mercado de trabajo de titulados universitarios. Madrid, Monthly Reports.

Shapley, L. and H. Scarf (1974). "On Cores and Indivisibility." Journal of Mathematical Economics, **1**: 23-37.

Sivey, P. M., A. Scott, J. Witt, J. Humphreys and C. Joyce (2010). "Why Junior Doctors Don't Want to Become General Practitioners: A Discrete Choice Experiment from the MABEL Longitudinal Study of Doctors." SSRN eLibrary.

StataCorp (2013). Stata 13 Base Reference Manual. College Station TX, Stata Press.

Steinbrook, R. (2009). "Easing the shortage in adult primary care—is it all about money?" New England Journal of Medicine **360**(26): 2696-2699.

Thistlethwaite, J. E., S. R. Leeder, M. R. Kidd and T. Shaw (2008). "Addressing general practice workforce shortages: policy options." Med J Aust **189**(2): 118-121.

Thornton, J. and F. Esposito (2003). "How important are economic factors in choice of medical specialty?" Health Econ **12**(1): 67-73.

Timmins, C. and J. Murdock (2007). "A revealed preference approach to the measurement of congestion in travel cost models." Journal of Environmental Economics and Management **53**: 230–249.

Train, K. (2009). Discrete Choice Methods With Simulation, 2nd Edition. New York, Cambridge University Press.

Watts, J. J. and L. Segal (2009). "Market failure, policy failure and other distortions in chronic disease markets." BMC Health Serv Res **9**: 102.

Table 1. Descriptive Statistics for the 2012 MIR Registry and the 2012 Post-MIR Survey

	<i>2012 MIR Registry</i>	<i>2012 Post-MIR Survey</i>	<i>P-value</i> <sup>a</sup>
<i>Sample size</i>	6,254	3,177	
<i>Spanish nationality</i> <sup>b</sup>	77.8%	77.3%	0.58
<i>Women</i> <sup>b</sup>	66.0%	67.1%	0.29
<i>MIR Ranking</i> <sup>c</sup>	3,307 (1,929)	3,049 (1,899)	0.00
<i>Distance from home to the training center (minutes)</i> <sup>c</sup>	153.3 (334.5)	138.9 (321.5)	0.05
<i>Home and training center in the same province</i> <sup>b</sup>	52.6%	56.4%	0.00

a. For binary variables (Spanish nationality, women, and home/training center in same province), P-value based two-group two tailed comparison test of proportions. For other variables (MIR ranking, distance from home to training center), P-value based on one-way ANOVA mean comparison test.

b. Mean values of binary variables.

c. Mean values with standard deviations in parentheses. The highest ranked candidate had a MIR ranking of 1.

Table 2. Preferred Specialties Reported in the 2012 Post-MIR Survey <sup>a</sup>

<i>Preferred Specialty</i>	<i>Number of Respondents</i>	<i>Assigned to Preferred Specialty (%)</i>		<i>Assigned to Family Medicine (%)<sup>c</sup></i>	
Cardiology	160	76	47.5%	18	11.2%
Plastic Surgery	110	13	11.8%	15	13.6%
Dermatology	193	50	25.9%	33	17.5%
FCM	234	229	97.9%	229	97.9%
Internal Medicine	147	105	71.4%	37	25.2%
Obstetrics & Gynecology	226	124	54.9%	60	26.6%
Pediatrics	396	231	58.3%	106	26.8%
Other Specialties	1,651	1,042	63.1%	278	16.8%
Total(b)	3,117	1,870	60.0%	776	24.9%

a. Based upon the question: “If you could choose specialty without regard to your score on the exam, what specialty would you have chosen?”

Table 3. Comparison of Selected Coefficients in Models II and III: Tests of Deviations from the Hypothesis of Independence of Irrelevant Alternatives <sup>a</sup>

<i>Explanatory Variable</i> <sup>b</sup>	<i>Model II</i> <sup>c</sup>	<i>Model III</i> <sup>d</sup>
Probability of obtaining employment ( $X_1$ ) × Ranking in MIR	-0.291 (0.216)	-0.505 (0.163)
Favorable lifestyle and work hours ( $X_2$ ) × Ranking in MIR	<b>0.192</b> (0.246)	<b>-0.465</b> (0.205)
Recognition by patients ( $X_3$ ) × Ranking in MIR	<b>0.322</b> (0.183)	<b>1.605</b> (0.150)
Prestige among colleagues ( $X_4$ ) × Ranking in MIR	<b>0.702</b> (0.347)	<b>-3.734</b> (0.299)
Opportunity for professional development ( $X_5$ ) × Ranking in MIR	-0.728 (0.313)	-1.479 (0.299)
Annual remuneration with 10–15 years experience ( $X_6$ ) × Ranking in MIR	<b>0.147</b> (0.032)	<b>-0.018</b> (0.026)
Proportion of compensation from private practice ( $X_7$ ) × Ranking in MIR	-1.973 (0.319)	-2.337 (0.230)

a. Complete results in Appendix C, Table C1. All estimated coefficients and standard errors (in parentheses) have been multiplied by  $10^4$ . Parameter estimates in boldface correspond to significant differences between Models II and III at the 5% level, based on Z-statistic:

$$Z = (b_{II} - b_{III}) / \sqrt{s_{II}^2 + s_{III}^2}, \text{ where } b_{II} \text{ and } b_{III} \text{ are the estimated parameters from the two models, while } s_{II} \text{ and } s_{III} \text{ are their respective standard errors.}$$

- b. The highest ranked candidate had a MIR ranking of  $R_1 = 1$ , while lower-ranked candidates had higher values of  $R_i$ .
- c. 2012 MIR Registry restricted to the 3,117 candidates who also responded to the 2012 Post-MIR Survey. The dependent variable was the actual training program elected by the candidate. Each candidate's choice set was restricted to those training programs still available when his turn came up in the MIR sequence.
- d. 2012 Post-MIR Survey of the same 3,117 candidates in Model II. The dependent variable was the candidate's preferred training program if he were ranked first. Each candidate's choice set was the entire set of all training programs.

Table 4. Welfare Analysis <sup>a</sup>

<i>Policy Scenario</i>	$\Delta V$	$\Delta \bar{U}$	$\Delta U$
<i>1 (Random Assignment)</i> <sup>b</sup>	370 [349, 393]	567 [528, 608]	-2235 [-2339, -2148]
<i>2A (Concentrated Reduction in Supply)</i>	146	-9720	-2881
<i>2B (Generalized Reduction in Supply)</i> <sup>b</sup>	100 [94, 105]	-9596 [-9685, -9518]	-2806 [-2852, -2769]
<i>3 (Upgrading Family and Community Medicine)</i> <sup>c</sup>	859	3192	3835

a. For definitions of alternative policy scenarios and the criterion functions  $\Delta V$ ,  $\Delta \bar{U}$ , and  $\Delta U$ , see Section 6. Simulation Strategies.

b. 95% credibility intervals in brackets.

c. The results for policy scenario 3 are for the specific values  $X_4 = 6.30$  on a 10-point scale and  $X_7 = 39.3$  on a percentage scale, which correspond to the valuations of these characteristics for all specialties other than family medicine.

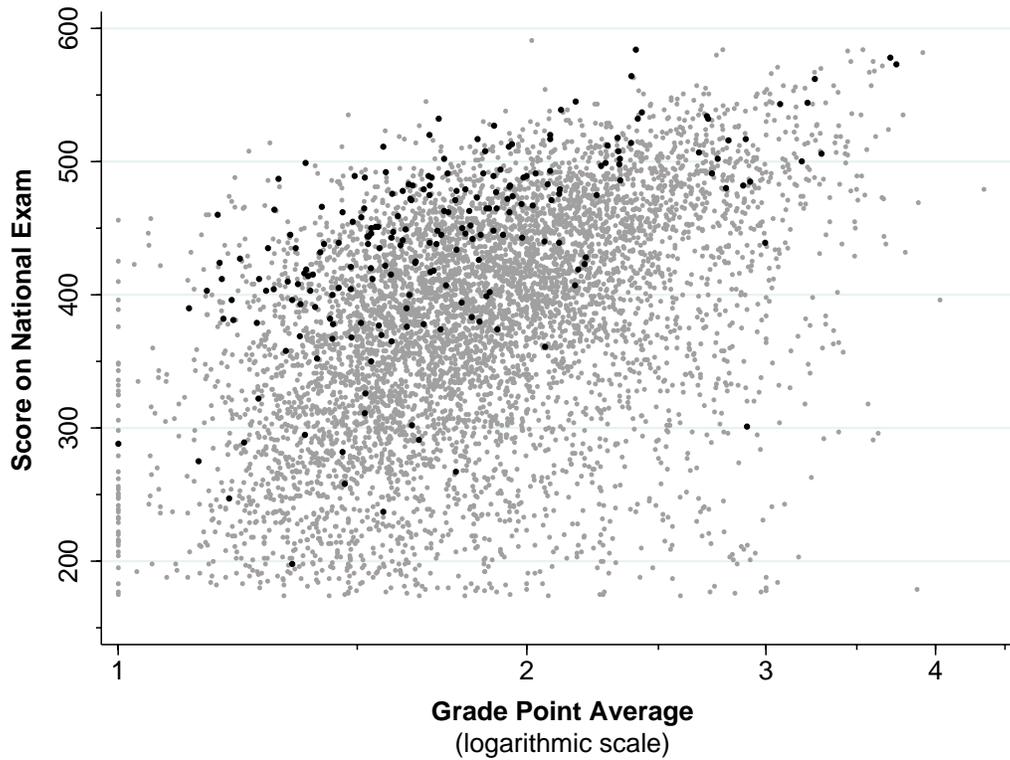


Figure 1. Scatterplot of each candidate's score on the national exam versus his medical school grade point average in the 2013 MIR cycle. Each point is one of 6,348 candidates. Black points represent 213 graduates of the medical school of the Universidad Autónoma de Madrid. The correlation coefficient between the national exam score and grade point average was +0.443.

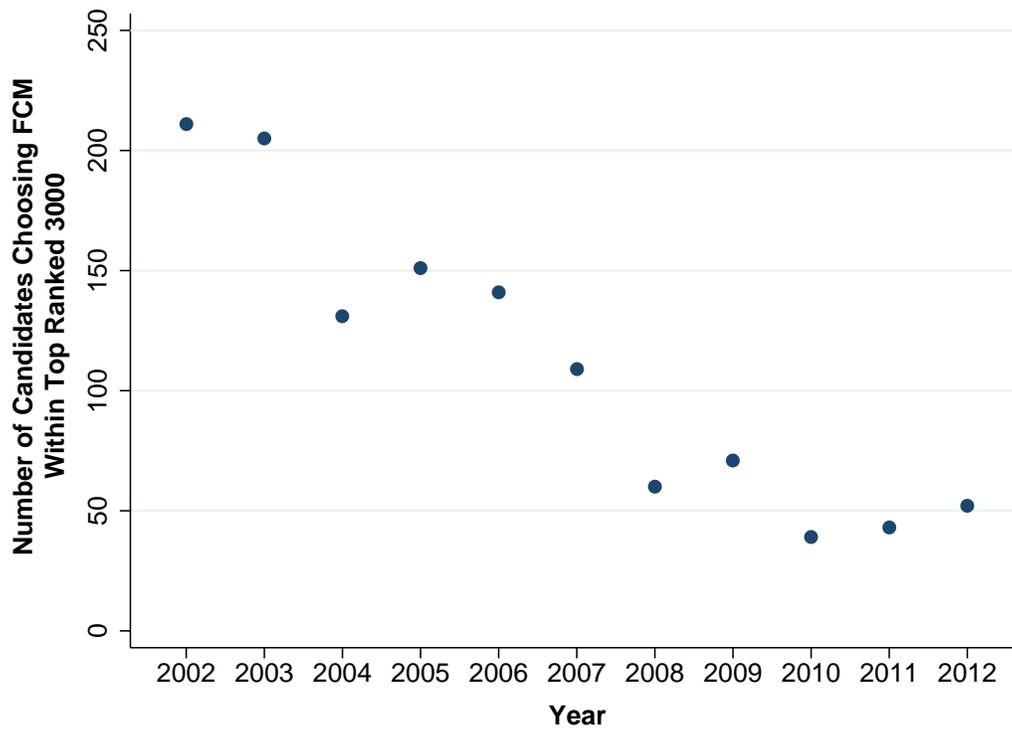


Figure 2. Number of candidates choosing family and community medicine (FCM) within the top ranked 3,000 candidates in each MIR cycle, 2002–2012. In each year, the total number of candidates exceeded 6,000.

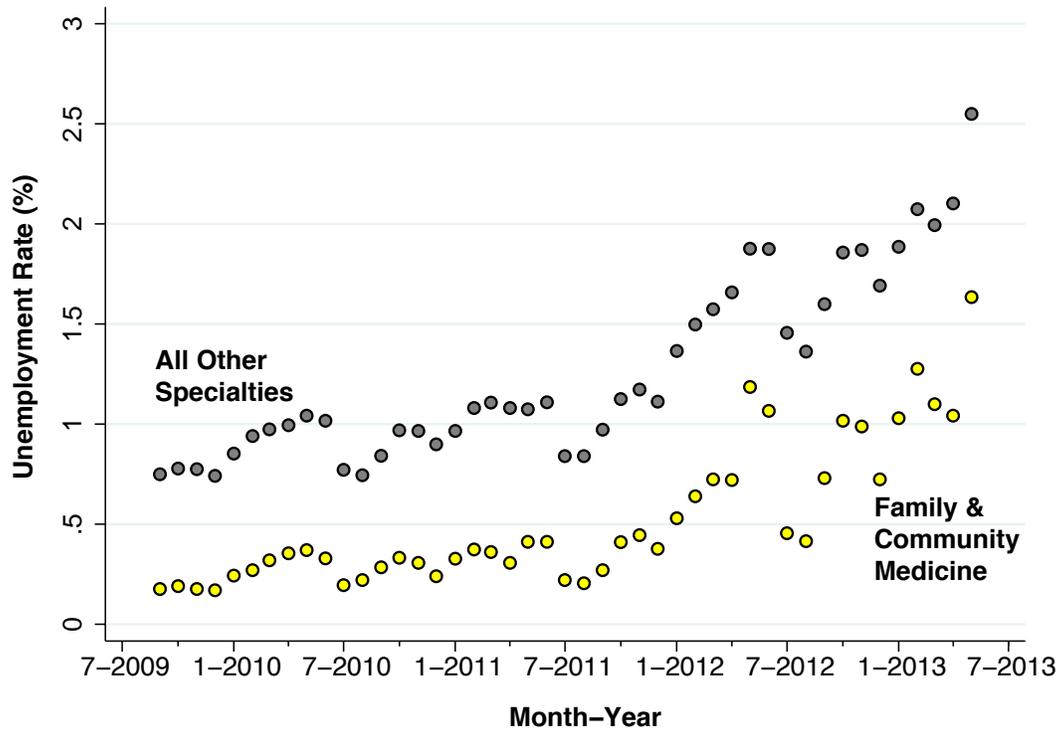


Figure 3. Monthly crude unemployment rate among Spanish physicians in family and community medicine (yellow points) and other specialties (dark gray filled points). No seasonal adjustment has been performed. Source: (Servicio Público de Empleo Estatal (SEPE) 2013)

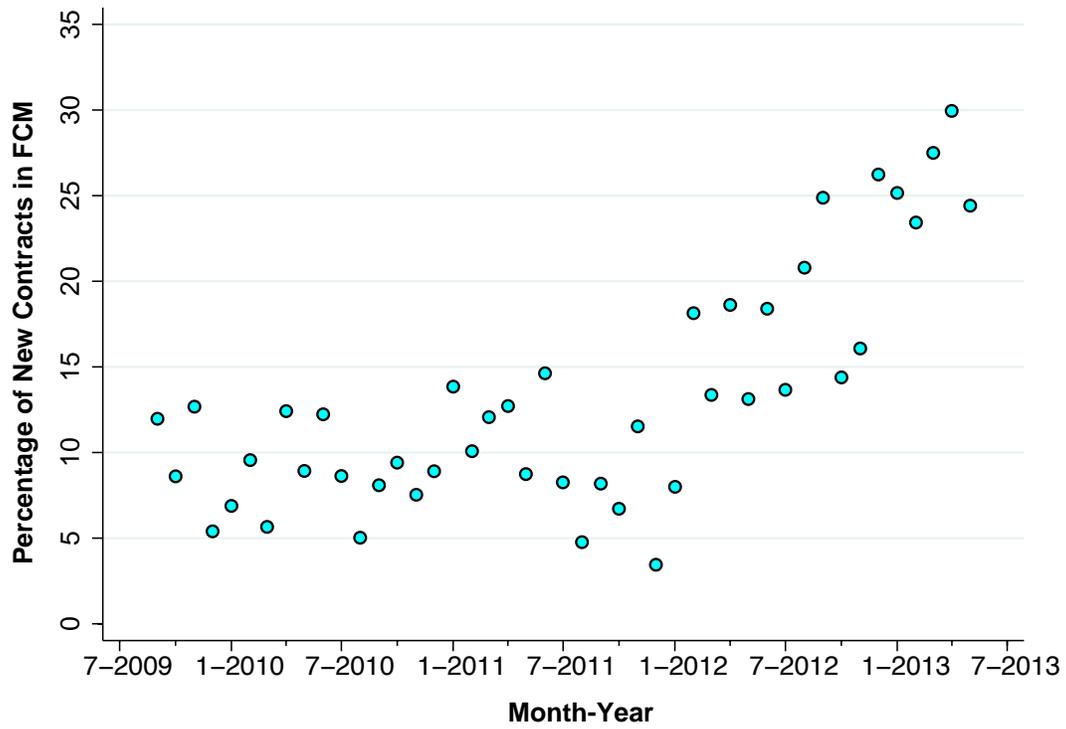


Figure 4. Monthly data on the percentage of new contracts in family and community medicine (FCM). Source: (Servicio Público de Empleo Estatal (SEPE) 2013)

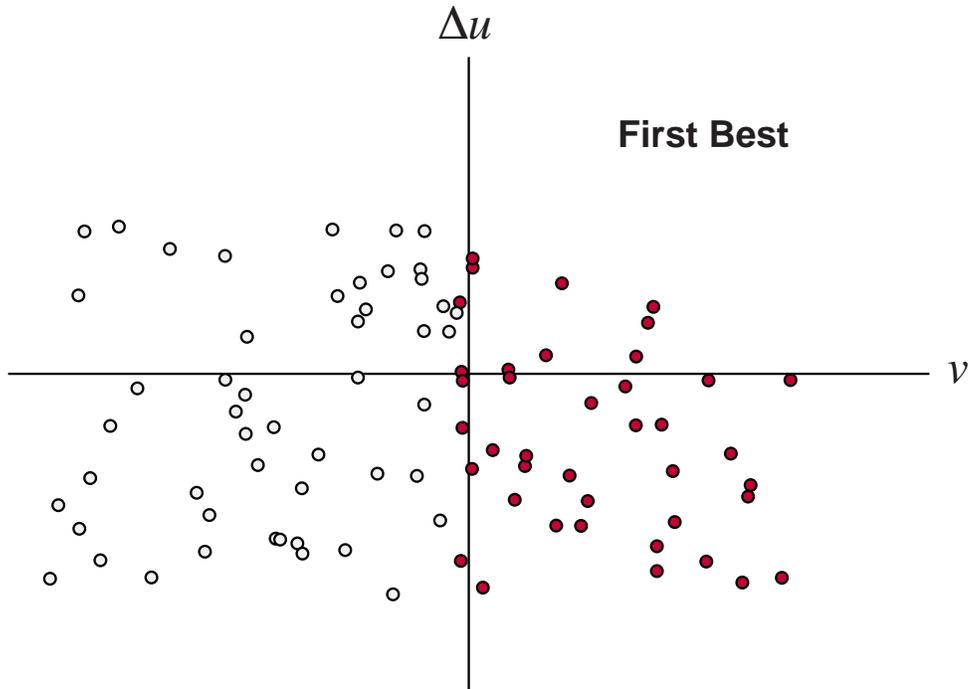


Figure 5. First best solution for a hypothetical set of  $N = 90$  candidates and  $M = 100$  training positions, of which  $M_F = 50$  are in family medicine and  $M_H = 50$  are in the other hospital-based specialty. Each candidate corresponds to a point in the  $(\nu, \Delta u)$  plane. At the first best optimum, a total of 40 candidates (red filled points) are assigned to family medicine, while the remaining 50 candidates (open points) are assigned to the hospital-based specialty.

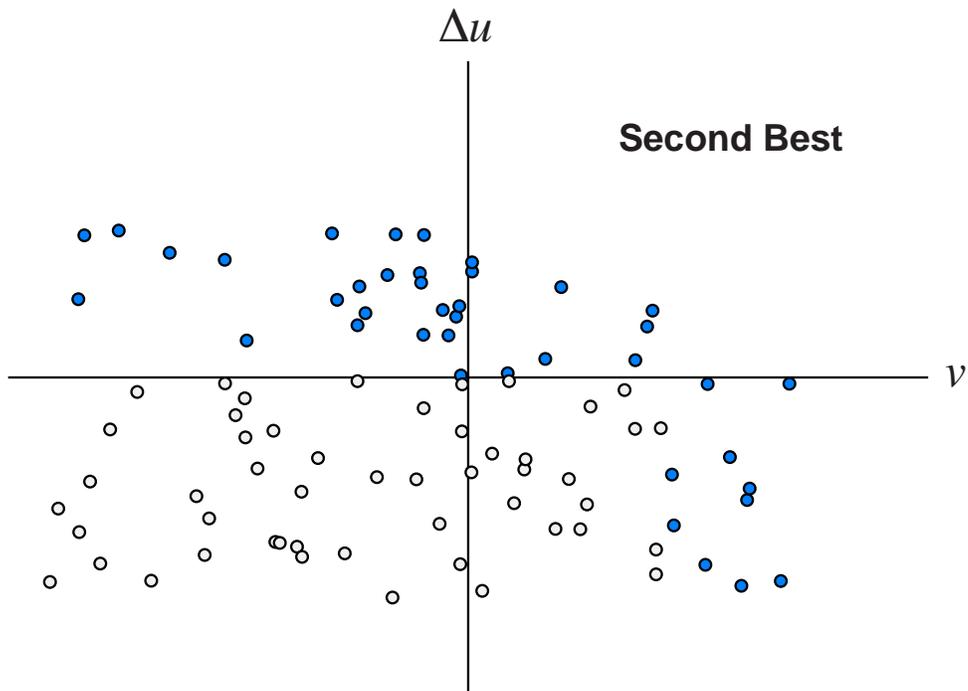


Figure 6. Second best solution among all priority mechanisms for the same hypothetical set of  $N = 90$  candidates and  $M = 100$  training positions shown in Figure 5. At the second best optimum, a total of 40 candidates (blue filled points) are assigned to family medicine, while the remaining 50 candidates (open points) are assigned to the hospital-based specialty.

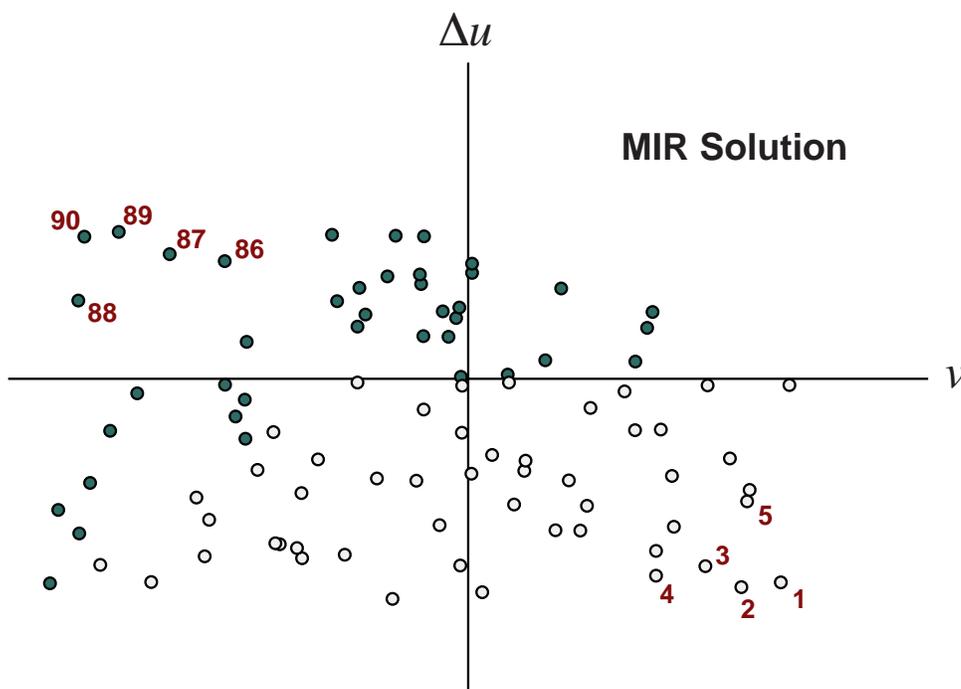


Figure 7. MIR solution for the same hypothetical set of  $N = 90$  candidates and  $M = 100$  training positions shown in Figures 5 and 6. A total of 40 candidates (green filled points) are assigned to family medicine, while the remaining 50 candidates (open points) are assigned to the hospital-based specialty. The five top-ranked candidates are labeled 1 through 5, while the five bottom-ranked candidates are labeled 86 through 90.

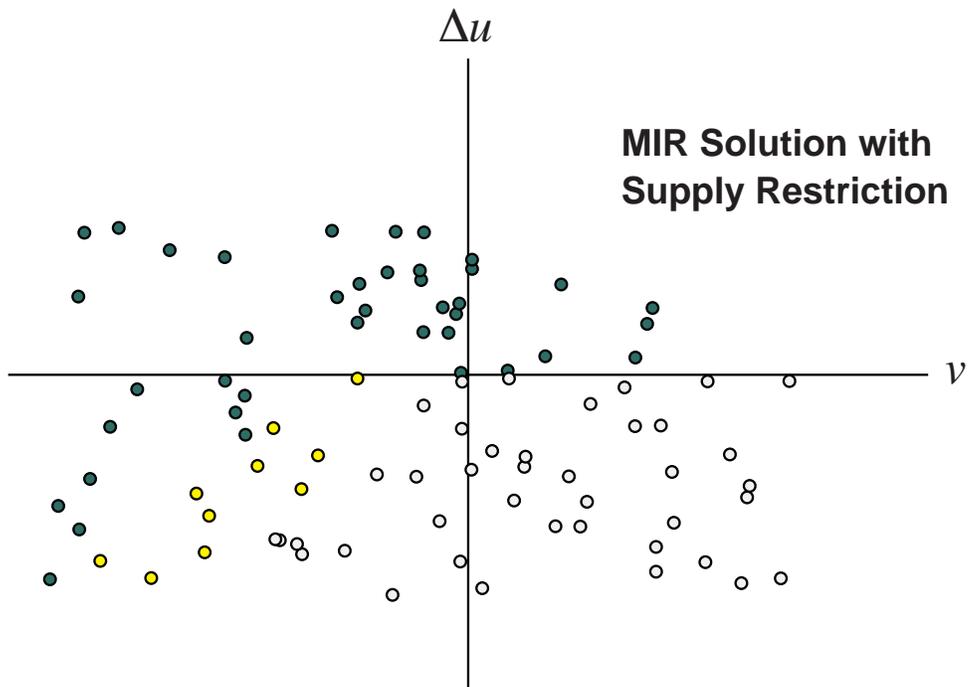


Figure 8. The MIR solution with a supply restriction. The green filled points show the candidates assigned to family medicine in the MIR solution of Figure 7. The 10 yellow points show the additional candidates assigned to family medicine if the number of positions in the hospital-based specialty is reduced from  $M_H = 50$  to  $M_H = 40$ .

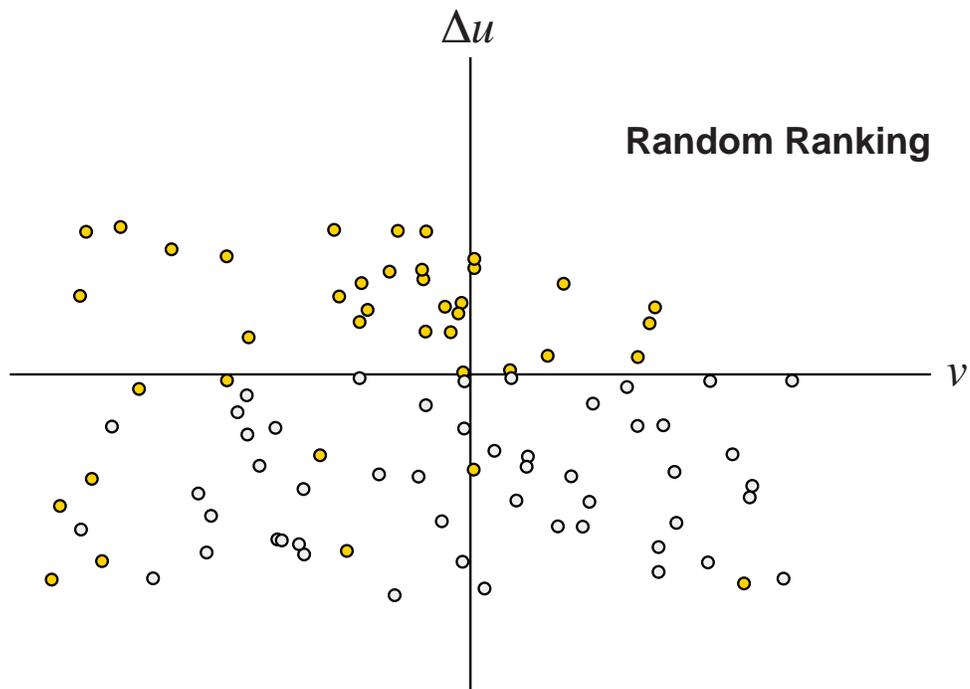


Figure 9. Priority mechanism with random ranking of candidates. The orange filled points show the 40 candidates assigned to family and community medicine.

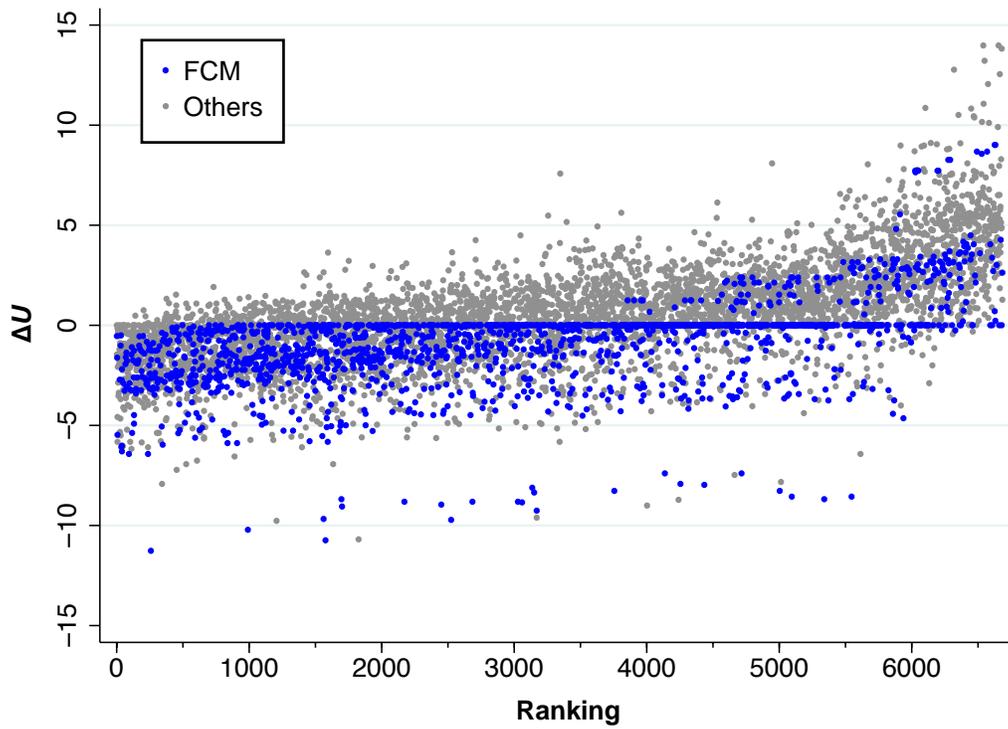


Figure 10. Change in Maximum Attained Utility  $\left(\Delta u_{i(t)} = u_{i(t)} - u_{i(0)}\right)$  versus Original MIR

Ranking  $\left(R_i\right)$  for the Random Assignment Policy Scenario (1). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

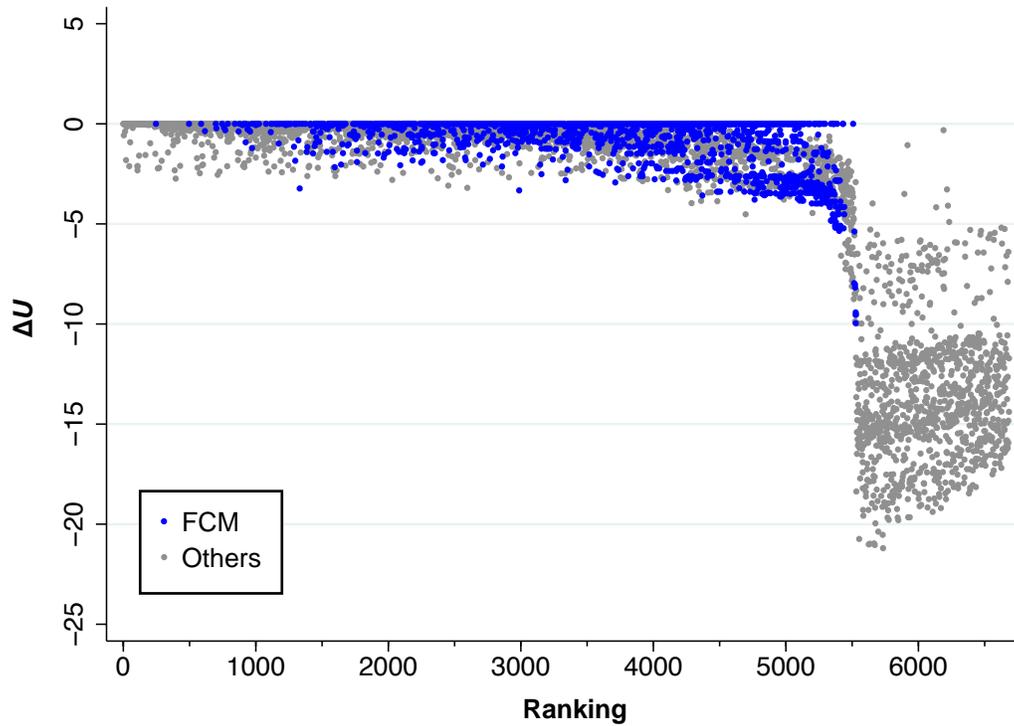


Figure 11. Change in Maximum Attained Utility ( $\Delta u_{i(2A)} = u_{i(2A)} - u_{i(0)}$ ) versus Original MIR Ranking ( $R_i$ ) for the Concentrated Reduction in Supply Policy Scenario (2A). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

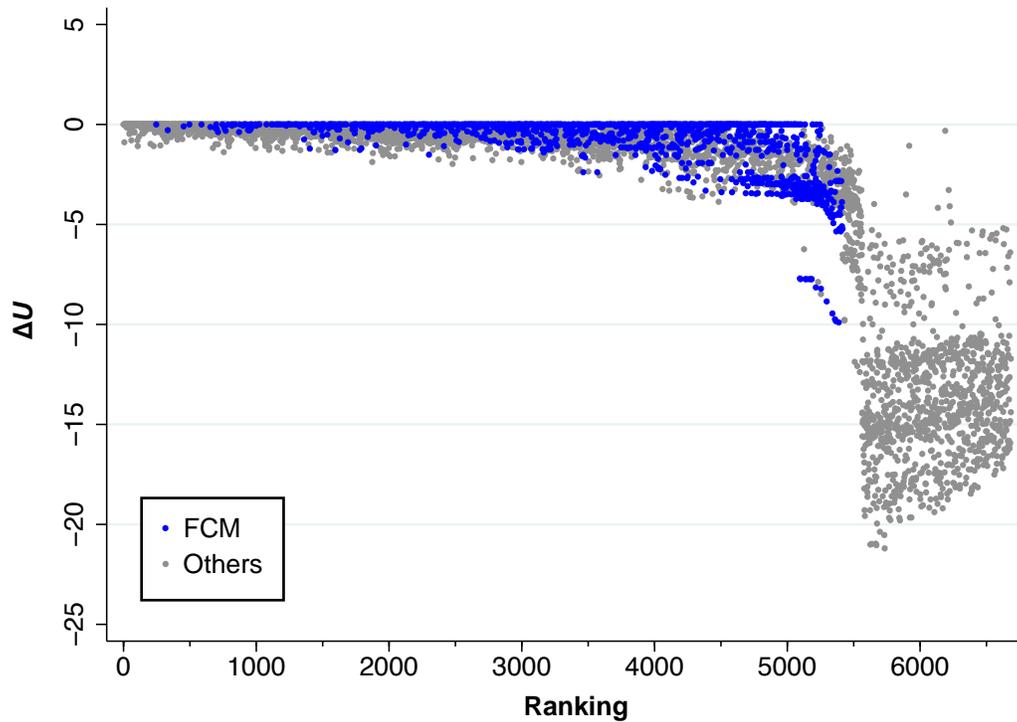


Figure 12. Change in Maximum Attained Utility ( $\Delta u_{i(2B)} = u_{i(2B)} - u_{i(0)}$ ) versus Original MIR Ranking ( $R_i$ ) for the Generalized Reduction in Supply Policy Scenario (2B). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

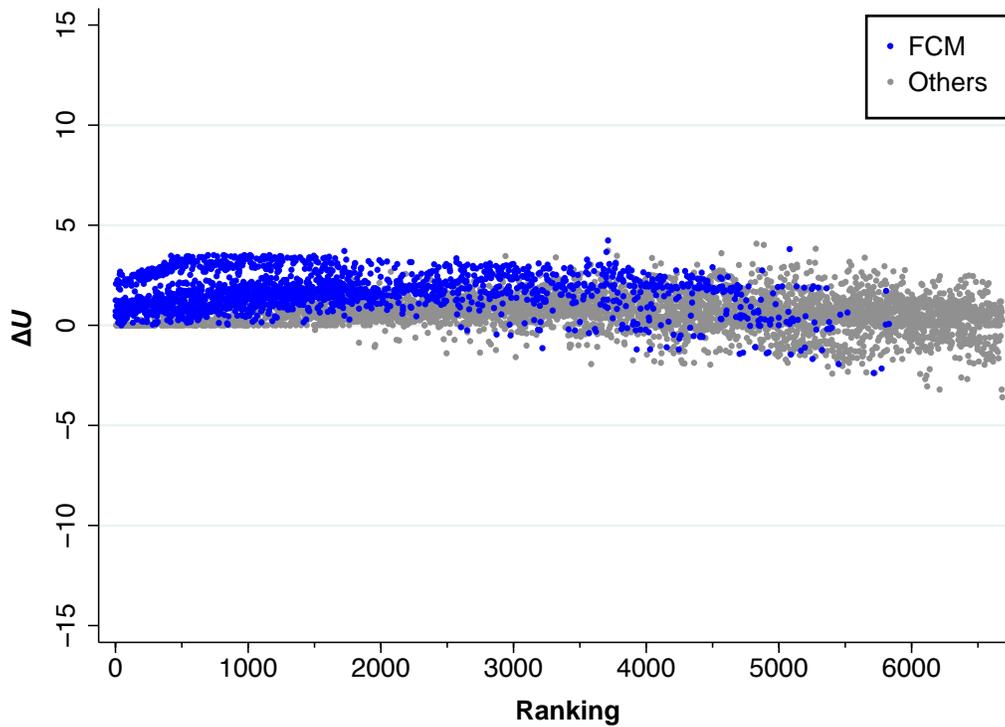


Figure 13. Change in Maximum Attained Utility ( $\Delta u_{i(3)} = u_{i(3)} - u_{i(0)}$ ) versus Original MIR

Ranking ( $R_i$ ) for the Upgrading Family and Community Medicine Policy Scenario (3). The blue points represent those candidates who were predicted to choose family and community medicine (FCM). The plot reflects an improvement in prestige among colleagues ( $X_4$ ) from 3.92 to 6.30 on a 10-point scale and an increase in the proportion of income from private practice ( $X_7$ ) from 13.2 to 39.3 on a percentage scale.

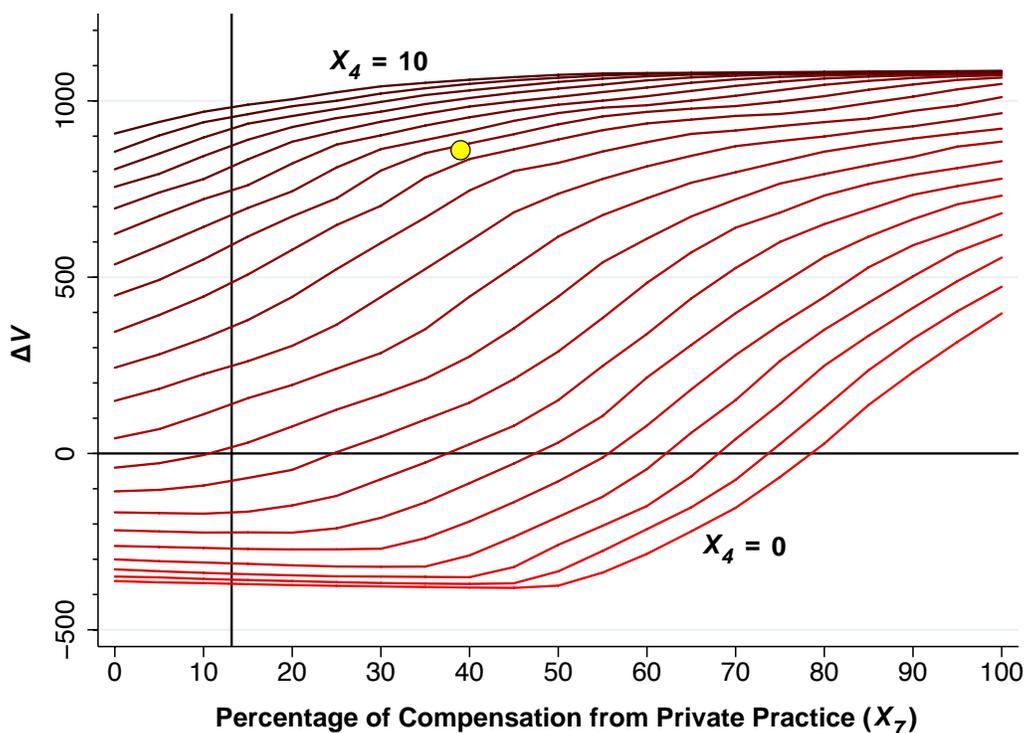


Figure 14. Simulations of Scenario 3 (Upgrading Family and Community Medicine) for Different Values of Prestige among Colleagues ( $X_4$ ) and Percentage of Compensation from Private Practice ( $X_7$ ). The vertical axis represents the change in the external benefit criterion  $\Delta V_{(3)} = V_{(3)} - V_{(0)}$ . The horizontal axis measures  $X_7$ , scaled in percentage terms from 0 to 100. The relative darkness of shade of each curve represents the level of  $X_4$ . The yellow point represents the specific values of  $X_4$  and  $X_7$  assumed in Table 5. The solid vertical line is set at  $X_7 = 13.2\%$ , which represents the value, in combination with  $X_4 = 3.92$ , at which the criterion  $\Delta V = 0$ .

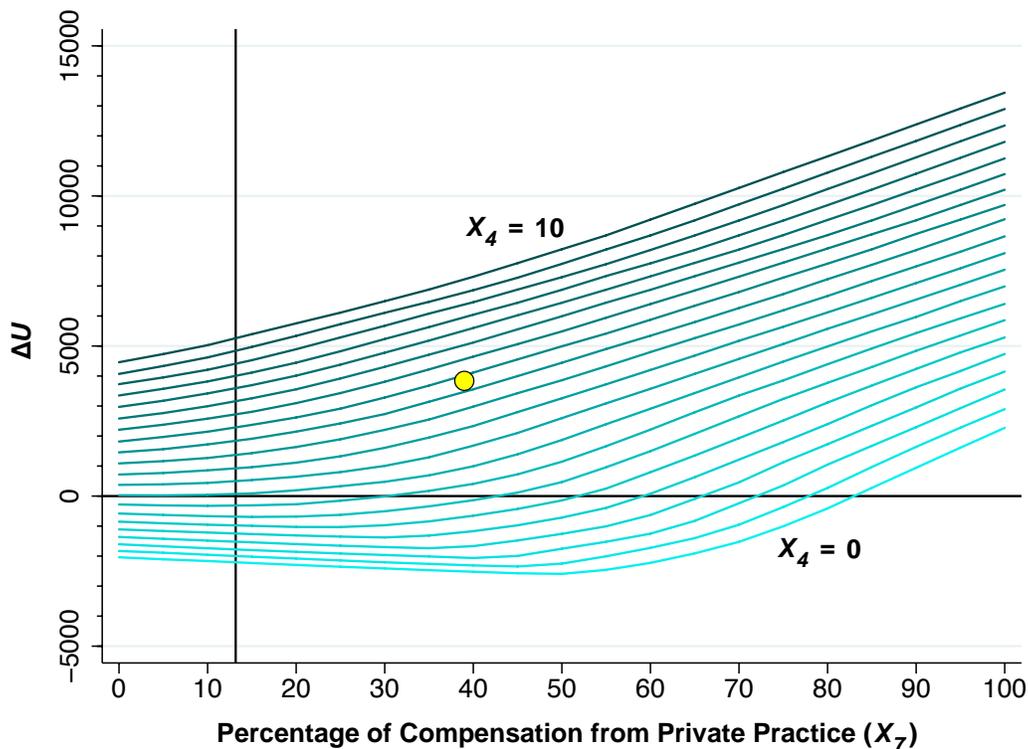


Figure 15. Simulations of Scenario 3 (Upgrading Family and Community Medicine) for Different Values of Prestige among Colleagues ( $X_4$ ) and Percentage of Compensation from Private Practice ( $X_7$ ). The vertical axis represents the change in the welfare criterion  $\Delta U_{(3)} = U_{(3)} - U_{(0)}$ . The horizontal axis measures  $X_7$ , scaled in percentage terms from 0 to 100.

The relative darkness of shade of each curve represents the level of  $X_4$ . The yellow point represents the specific values of  $X_4$  and  $X_7$  assumed in Table 5. The solid vertical line is set at  $X_7 = 13.2\%$ , which represents the value, in combination with  $X_4 = 3.92$ , at which the criterion  $\Delta U = 0$ .

## Appendix A. Optimal and Second-Best Assignment Mechanisms with Supply Restrictions

In the main text, we assumed that the number of positions equaled or exceeded the number of candidates, that is,  $N \leq M$ . Here, we consider supply restrictions that result in fewer positions than candidates, that is,  $N > M$ .

If there are fewer positions than candidates, then  $N - M > 0$  candidates will have to end up with a null position ( $O$ ). Let  $O$  denote a set of  $N - M$  such null positions. Then an assignment becomes a mapping between  $C$  and  $P \cup O$ .

We now let an assignment map each candidate  $c_i$  into an ordered pair of binary numbers  $(f_{i1}, f_{i2})$ . When  $(f_{i1}, f_{i2}) = (1, 0)$ , the candidate is assigned to family and community medicine. When  $(f_{i1}, f_{i2}) = (0, 0)$ , the candidate is assigned to the hospital-based specialty, and when  $(f_{i1}, f_{i2}) = (0, 1)$ , he is assigned to the null specialty. No candidate receives the assignment  $(f_{i1}, f_{i2}) = (1, 1)$ . In effect,  $f_i$  is a two-bit indicator where the first bit  $f_{i1}$  indicates whether the candidate was assigned to family and community medicine. The second bit  $f_{i2}$  indicates whether a candidate who has not been assigned to family and community has instead by assigned to the null specialty.

A candidate's attained utility is now  $u_i = f_{i1}u_{iF} + f_{i2}u_{iO} + (1 - f_{i1} - f_{i2})u_{iH}$ , where  $u_{iO}$  is the utility attained from the null specialty. It is natural to assume that every candidate would prefer some specialty assignment over the null specialty, that is,  $u_{iF} > u_{iO}$  and  $u_{iH} > u_{iO}$ .<sup>8</sup> We can rewrite this attained utility as  $u_i = f_{i1}\Delta u_{i1} + f_{i2}\Delta u_{i2} + u_{iH}$ , where  $\Delta u_{i1} = u_{iF} - u_{iH}$  and  $\Delta u_{i2} = u_{iO} - u_{iH}$ , and where we have assumed that  $\Delta u_{i2} < 0$  and  $\Delta u_{i1} > \Delta u_{i2}$ . The social welfare function can then be formulated as

$$(A1) \quad W = \sum_{i=1}^N (f_{i1}v_{i1} + f_{i2}v_{i2}) + W_0,$$

---

<sup>8</sup> In reality, some candidates withdraw from the MIR competition when there are still available training positions. Most plan to retake the national exam the following year in the hope of attaining a higher ranking.

where  $v_{i1} = r_i + \lambda w_i \Delta u_{i1}$ ,  $v_{i2} = \lambda w_i \Delta u_{i2}$ , and  $W_0$  is defined in (5).

In (A1), the contribution to social welfare of a candidate assigned to family and community medicine is still  $v_{i1}$ . But if the candidate is not assigned to that specialty and is instead assigned to the null specialty, then his contribution is  $v_{i2}$ , which is necessarily negative as the result of the loss of utility for having been excluded.

It may appear that the extension of the set of positions from  $\mathbf{P}$  to  $\mathbf{P} \cup \mathbf{O}$  constitutes nothing more than the addition of a third specialty. But this is not exactly the case. We have assumed that no candidate would choose the null specialty when his turn comes up, that is,  $u_{iF} > u_{iO}$  and  $u_{iH} > u_{iO}$ . Accordingly, in a priority mechanism, the only candidates who end up with the null specialty are those at the bottom of the ranking when positions in the two other specialties ( $F$  and  $H$ ) have already run out.

Consider a priority mechanism. In the first step (2.1), we rank candidates with  $v_{i1} > 0$  and  $\Delta u_{i1} \geq 0$  at the top of the ordering, and they elect family and community medicine, just as before. Assuming that this step does not exhaust all positions in family and community medicine, we continue with the second step (2.2), ranking candidates with  $\Delta u_{i1} < 0$  and the algebraically lowest values of  $v_{i1}$ . Let's assume that this step exhausts all hospital-based positions. What remains is a set of candidates with values of  $v_{i1}$  and  $v_{i2}$ , all of whom will have to elect family and community medicine or end up excluded. In step 2.3, we now rank the remaining candidates in decreasing order of  $v_{i1} - v_{i2} = r_i + \lambda w_i (u_{iF} - u_{iO})$ . That's because the effect on social welfare of moving a candidate from a null position to a position in family and community medicine is  $v_{i1} - v_{i2}$ . If step 2.2 does not exhaust all hospital-based positions, then we have left a subset of candidates with negative values of  $v_{i1}$  and positive values of  $\Delta u_{i1}$ . In step 2.3, we rank these candidates in decreasing order of  $v_{i1}$  until the set of positions in family and community medicine runs out. Then, with only positions in the hospital-based specialty or the null specialty available, we rank the remaining candidates in increasing order of  $v_{i2}$ , which is the effect on social welfare of moving a candidate from the null position to a hospital-based position.

**Appendix B. Attributes of Medical Specialties**Table B1. Estimated Valuations of Seven Attributes for Each of 44 Medical Specialties <sup>a,b,c</sup>

<i>Specialty</i>	<i>X<sub>1</sub>. Probability of Obtaining Employment</i>	<i>X<sub>2</sub>. Lifestyle &amp; Work Hours</i>	<i>X<sub>3</sub>. Recognition by Patients</i>	<i>X<sub>4</sub>. Prestige among Colleagues</i>	<i>X<sub>5</sub>. Opportunity for Professional Development</i>	<i>X<sub>6</sub>. Annual Remuneration with 10–15 Years Experience</i>	<i>X<sub>7</sub>. Proportion of Compensation from Private Practice</i>
Allergy	5.48	8.28	5.50	4.74	6.47	73	4.42
Clinical Analysis & Laboratory Medicine	5.59	8.13	2.38	3.29	5.74	71	3.83
Anatomy & Pathology	5.86	8.46	2.41	4.00	6.12	67	2.67
Anesthesiology	7.92	6.32	4.29	6.48	6.83	93	3.89
Vascular Surgery	6.42	5.25	8.17	8.69	8.42	103	4.52
Gastroenterology	6.68	6.51	7.08	6.69	7.25	83	3.87
Biochemistry	4.88	8.14	1.77	2.97	6.15	61	2.77
Cardiology	7.18	5.86	8.65	8.67	8.17	97	4.49
Cardiovascular Surgery	6.78	5.00	9.13	9.04	8.32	104	4.35
General Surgery	6.88	4.85	8.27	7.79	7.78	93	3.85
Oral & Maxillofacial Surgery	6.37	5.44	7.95	7.57	7.27	114	5.68
Traumatology	7.57	5.52	7.59	6.88	7.93	95	4.65
Pediatric Surgery	6.60	5.49	8.60	8.37	8.12	96	4.14
Plastic Surgery	7.45	6.99	8.63	8.05	8.50	135	7.07
Thoracic Surgery	6.03	5.87	8.11	8.08	7.48	100	4.29
Dermatology	7.63	9.13	7.05	7.38	7.38	103	5.89
Endocrinology	7.16	8.09	7.07	6.84	7.69	81	4.51
Pharmacology	4.51	8.25	2.48	2.79	6.30	69	1.87
Geriatrics	7.53	7.83	6.23	4.24	5.67	72	3.17
Hematology	5.64	6.84	6.44	6.24	7.46	79	2.84
Immunology	4.23	7.84	3.28	4.07	8.27	69	2.45
Occupational Medicine	5.21	8.34	4.23	3.57	4.03	62	5.09

<i>Specialty</i>	<i>X<sub>1</sub>. Probability of Obtaining Employment</i>	<i>X<sub>2</sub>. Lifestyle &amp; Work Hours</i>	<i>X<sub>3</sub>. Recognition by Patients</i>	<i>X<sub>4</sub>. Prestige among Colleagues</i>	<i>X<sub>5</sub>. Opportunity for Professional Development</i>	<i>X<sub>6</sub>. Annual Remuneration with 10–15 Years Experience</i>	<i>X<sub>7</sub>. Proportion of Compensation from Private Practice</i>
Family & Community Medicine	8.39	7.76	5.92	3.92	5.11	56	1.32
Rehabilitation Medicine	5.77	8.49	6.52	4.72	6.28	73	4.42
Intensive Care Medicine	6.49	4.74	7.11	7.04	7.15	82	2.31
Internal Medicine	6.64	5.64	6.72	6.59	6.43	80	2.92
Nuclear Medicine	6.50	7.91	4.25	4.56	7.54	82	3.57
Preventive Medicine & Public Health	5.60	8.79	3.35	4.19	6.47	63	2.09
Microbiology	5.46	8.27	2.27	3.18	5.83	70	1.88
Nephrology	5.72	6.82	6.49	5.91	6.59	80	2.94
Pulmonology	6.04	7.19	6.68	6.16	6.59	82	3.35
Neurosurgery	6.62	5.03	8.48	8.70	8.89	100	4.11
Clinical Neurophysiology	4.23	7.18	3.71	3.83	6.60	77	2.59
Neurosurgery	6.08	6.42	7.36	7.73	8.27	92	4.06
Obstetrics & Gynecology	7.23	4.95	7.09	6.82	7.31	96	4.92
Ophthalmology	6.76	7.81	7.25	6.90	7.29	108	6.03
Medical Oncology	7.07	6.53	7.99	7.30	8.76	83	3.33
Radiotherapy Oncology	6.56	6.79	6.84	6.39	8.27	80	2.46
Otolaryngology	6.89	6.86	7.32	7.11	7.80	87	4.94
Pediatrics	7.97	6.67	7.65	7.20	7.40	80	4.21
Psychiatry	6.42	7.10	5.49	5.01	6.00	85	4.37
Radiology	7.34	7.75	3.36	5.26	7.38	90	3.99
Rheumatology	5.54	7.20	5.80	5.64	6.79	76	3.34
Urology	6.36	6.63	6.77	6.47	7.00	92	4.13

## Notes to Table B1:

- a. Source: 2011 survey to 978 final-year students in Spanish medical schools (Harris, Gonzalez Lopez-Valcarcel et al. 2013). The introductory survey text was, “In this section, you’ll define your profile of some medical specialties, including the one that you’ve just chosen as your first choice as well as others chosen at random. Think about your *perceptions and expectations* concerning each specialty.”. The preamble to the two questions on attributes 6 and 7 was: “The following questions are about compensation. To facilitate your responses, recall that the average annual gross income of a full-time specialist in Family & Community Medicine with 10–15 years experience is currently about 60,000 euros.”
- b. The survey questions concerning each specific attribute were as follows:
- X1. *Probability of Obtaining Employment*. “How would you rate the probability of obtaining work in the next three years, whether in the public or private sector, for an individual who became certified in this specialty today? (0 to 100 percent)” It has been rescaled to 0-10
- X2. *Lifestyle & Work Hours*. “Work hours, working conditions, and the ability to reconcile work with family life. (0 to 10, 0 = very bad, 10 = very good)”
- X3. *Recognition by Patients*. “Recognition of professional work on the part of patients. (0 to 10)”
- X4. *Prestige among Colleagues*. “Prestige and recognition among colleagues as well as social recognition. (0 to 10)”
- X5. *Opportunity for Professional Development*. “Possibility of promotion or future professional development within the specialty (new fields, new techniques, scientific advances). (0 to 10)”
- X6. *Annual Remuneration with 10–15 Years Experience*. “Estimate the current average annual gross remuneration (public and private combined) of a specialist with 10–15 years of experience. (Thousands of euros)”
- X7. *Proportion of Compensation from Private Practice*. “What percentage of this remuneration (including public and private) do you believe comes from private practice? (0 to 100 percent)” It has been rescaled to 0-10.
- c. The valuations corresponded to the predicted values from the following model:  $Y_{ijk} = \mu + \theta_{jk} + v_i + \varepsilon_{ijk}$ , where  $Y_{ijk}$  is the reported assessment by medical student  $i$  of attribute  $j$  for specialty  $k$ ,  $\mu$  and  $\theta_{jk}$  are unknown parameters, and  $v_i$  and  $\varepsilon_{ijk}$  are error terms.

**Appendix C. Detailed Model Results**

Table C1. Parameter Estimates for Models I, II and III

	<i>Model I</i> <sup>a</sup>	<i>Model II</i> <sup>b</sup>	<i>Model III</i> <sup>c</sup>
Mean ( $\mu$ )			
Distance (minutes) from home province to training center	-0.0031 *** (0.0001)	-0.0029 *** (0.0002)	-0.0024 *** (0.0002)
PET scan available in training center hospital	0.040 (0.035)	0.054 (0.048)	-0.011 (0.050)
Training center in same province as candidate's residence	2.601 *** (0.048)	2.759 *** (0.069)	2.721 *** (0.065)
Number of hospital beds in training center	0.00007 (0.00004)	0.00007 (0.00006)	-0.00009 * (0.00005)
Number of positions available in training program	0.0118 *** (0.0003)	0.0117 *** (0.0004)	0.0074 *** (0.0004)
Favorable lifestyle and work hours ( $X_2$ )	0.214 *** (0.074)	0.209 ** (0.100)	-0.040 (0.083)
Prestige among colleagues ( $X_4$ )	0.978 *** (0.115)	1.022 *** (0.158)	1.267 *** (0.139)
Opportunity for professional development ( $X_5$ )	0.236 ** (0.100)	0.131 (0.136)	0.185 (0.138)
Annual Remuneration with 10–15 years experience ( $X_6$ )	-0.018 * (0.010)	-0.016 (0.014)	-0.016 (0.011)
Probability of obtaining employment ( $X_7$ ) $\times$ Female gender	0.066 (0.046)	0.058 (0.064)	0.033 (0.049)
Favorable lifestyle and work hours ( $X_2$ ) $\times$ Female gender	0.071 (0.048)	0.171 *** (0.065)	0.131 ** (0.056)
Recognition by patients ( $X_3$ ) $\times$ Female gender	-0.042 (0.037)	-0.054 (0.050)	-0.044 (0.046)
Prestige among colleagues ( $X_4$ ) $\times$ Female gender	0.265 *** (0.072)	0.314 *** (0.099)	0.247 *** (0.093)
Opportunity for professional development ( $X_5$ ) $\times$ Female gender	-0.139 (0.064)	-0.097 (0.088)	-0.033 (0.092)

	<i>Model I</i> <sup>a</sup>	<i>Model II</i> <sup>b</sup>	<i>Model III</i> <sup>c</sup>
Annual Remuneration with 10–15 years experience ( $X_6$ ) × Female gender	-0.041 *** (0.006)	-0.037 *** (0.088)	-0.033 *** (0.008)
Proportion of compensation from private practice ( $X_7$ ) × Female gender	0.223 *** (0.060)	0.163 ** (0.084)	0.152 ** (0.072)
Probability of obtaining employment ( $X_7$ ) × Spanish nationality	0.364 *** (0.055)	0.326 *** (0.075)	0.119 ** (0.057)
Favorable lifestyle and work hours ( $X_2$ ) × Spanish nationality	-0.281 *** (0.062)	-0.333 *** (0.084)	-0.271 *** (0.069)
Recognition by patients ( $X_3$ ) × Spanish nationality	0.265 *** (0.044)	0.202 *** (0.060)	0.093 *** (0.054)
Prestige among colleagues ( $X_4$ ) × Spanish nationality	-0.617 *** (0.091)	-0.639 *** (0.125)	-0.328 *** (0.112)
Opportunity for professional development ( $X_5$ ) × Spanish nationality	-0.1802 ** (0.0750)	-0.1578 * (0.1010)	0.0002 (0.1074)
Annual Remuneration with 10–15 years experience ( $X_6$ ) × Spanish nationality	0.003 (0.008)	-0.005 (0.011)	-0.015 * (0.009)
Proportion of compensation from private practice ( $X_7$ ) × Spanish nationality	-0.248 *** (0.072)	-0.206 ** (0.100)	0.071 (0.084)
Probability of obtaining employment ( $X_7$ ) × Ranking in MIR <sup>d</sup>	-0.00005 *** (0.00002)	-0.00003 (0.00002)	-0.00005 *** (0.00002)
Favorable lifestyle and work hours ( $X_2$ ) × Ranking in MIR <sup>d</sup>	0.00003 * (0.00002)	0.00002 (0.00003)	-0.00005 ** (0.00002)
Recognition by patients ( $X_3$ ) × Ranking in MIR <sup>d</sup>	0.00002 * (0.00001)	0.00003 (0.00002)	0.00016 *** (0.00002)
Prestige among Colleagues ( $X_4$ ) × Ranking in MIR <sup>d</sup>	0.00010 *** (0.00003)	0.00007 ** (0.00003)	-0.00037 *** (0.00003)
Opportunity for professional development ( $X_5$ ) × Ranking in MIR <sup>d</sup>	-0.00012 *** (0.00002)	-0.00007 ** (0.00003)	-0.00015 *** (0.00003)
Annual Remuneration with 10–15 years experience ( $X_6$ ) × Ranking in MIR <sup>d</sup>	0.000019 *** (0.000002)	0.000015 *** (0.000003)	0.000002 (0.000003)

	<i>Model I</i> <sup>a</sup>	<i>Model II</i> <sup>b</sup>	<i>Model III</i> <sup>c</sup>
Proportion of compensation from private practice ( $X_7$ ) $\times$ Ranking in MIR <sup>d</sup>	-0.00022 *** (0.00002)	-0.00020 *** (0.00003)	-0.00023 *** (0.00002)
Probability of obtaining employment ( $X_1$ )	0.510 *** (0.074)	0.505 *** (0.100)	0.438 *** (0.079)
Recognition by patients ( $X_3$ )	-0.247 *** (0.065)	-0.242 *** (0.089)	-0.609 *** (0.071)
Proportion of compensation from private practice ( $X_7$ )	0.919 *** (0.110)	0.932 *** (0.148)	0.819 *** (0.116)
Standard Deviation ( $\hat{\sigma}$ )			
Probability of obtaining employment ( $X_1$ )	0.703 *** (0.044)	0.664 *** (0.061)	0.0003 (0.074)
Recognition by patients ( $X_3$ )	0.278 *** (0.035)	0.246 *** (0.053)	0.0003 (0.027)
Proportion of compensation from private practice ( $X_7$ )	0.416 *** (0.051)	0.396 *** (0.076)	0.002 (0.056)
Total number of choices	6,884,904	3,775,333	7,876,659
Number of candidates	6,254	3,117	3,117

- a. 2012 MIR Registry for all 6,254 candidates. The dependent variable was the actual training program elected by the candidate. Each candidate's choice set was restricted to those training programs still available when his turn came up in the MIR sequence.
- b. 2012 MIR Registry restricted to the 3,117 candidates who also responded to the 2012 Post-MIR Survey. The dependent variable was the actual training program elected by the candidate. Each candidate's choice set was restricted to those training programs still available when his turn came up in the MIR sequence.
- c. 2012 Post-MIR Survey of the same 3,117 candidates in Model II. The dependent variable was the candidate's preferred training program if he were ranked first. Each candidate's choice set was the entire set of all training programs.
- d. The highest ranked candidate had a MIR ranking of  $R_1 = 1$ , while lower ranked candidates had higher values of  $R_i$ .

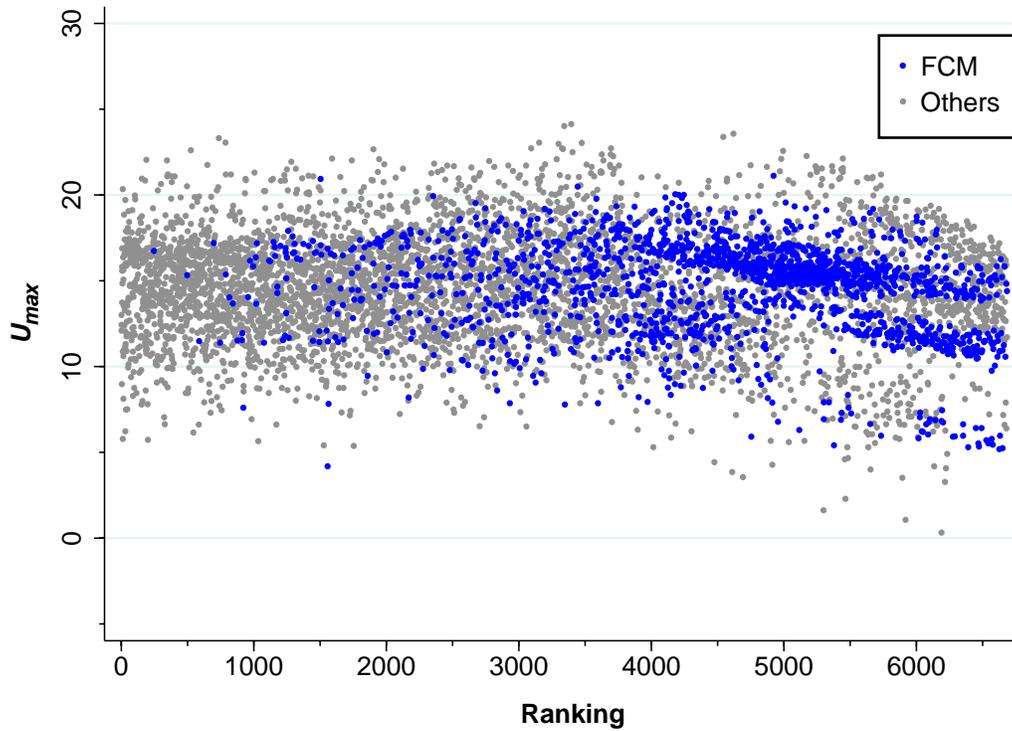
**Appendix D. Plots of Maximum Attained Utility versus Original MIR Ranking for each Policy Scenario**

Figure D1. Maximum Attained Utility ( $u_i$ ) versus MIR Ranking ( $R_i$ ) in the Baseline Scenario (0). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

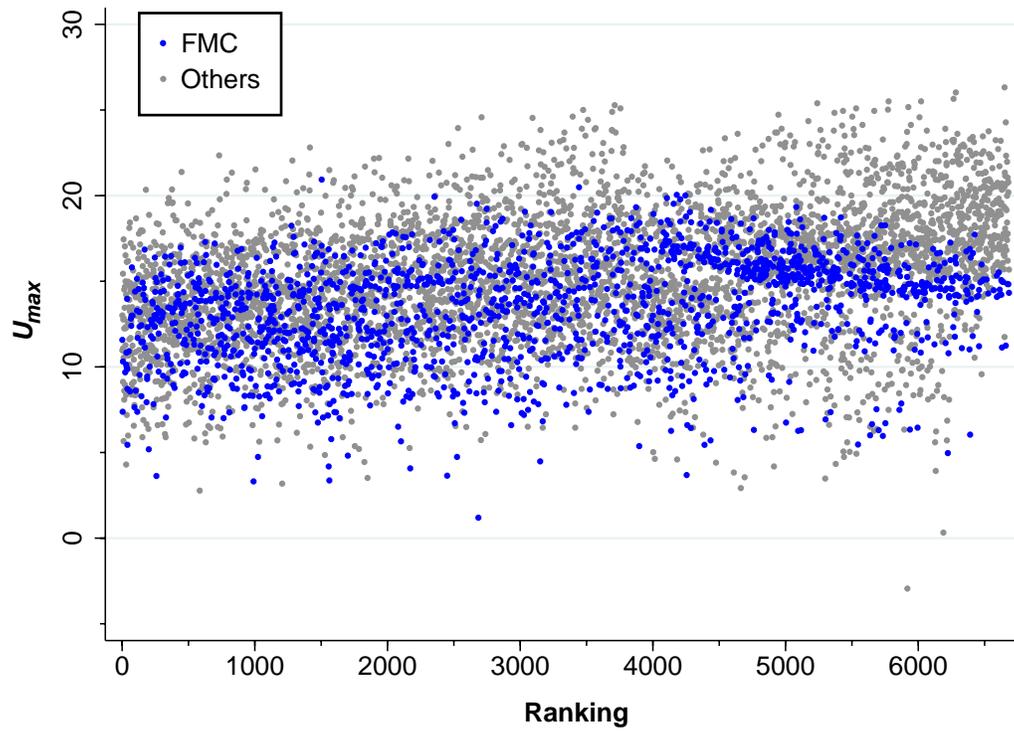


Figure D2. Maximum Attained Utility ( $u_i$ ) versus MIR Ranking ( $R_i$ ) in the Random Assignment Policy Scenario (1). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

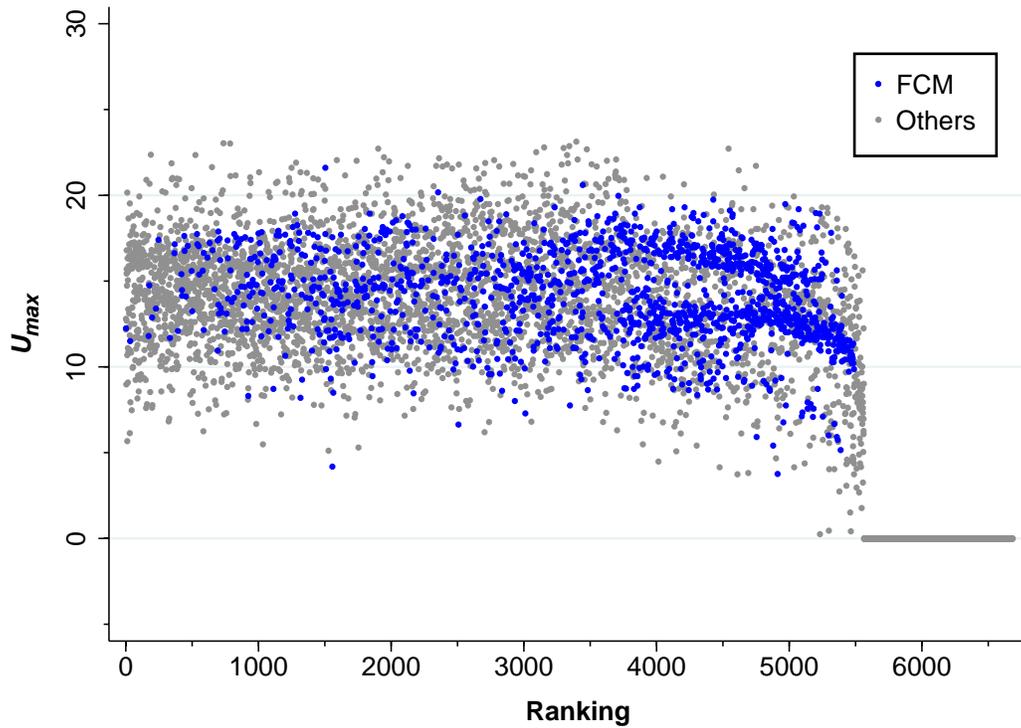


Figure D3. Maximum Attained Utility ( $u_i$ ) versus MIR Ranking ( $R_i$ ) in the Concentrated Reduction in Supply Policy Scenario (2A). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

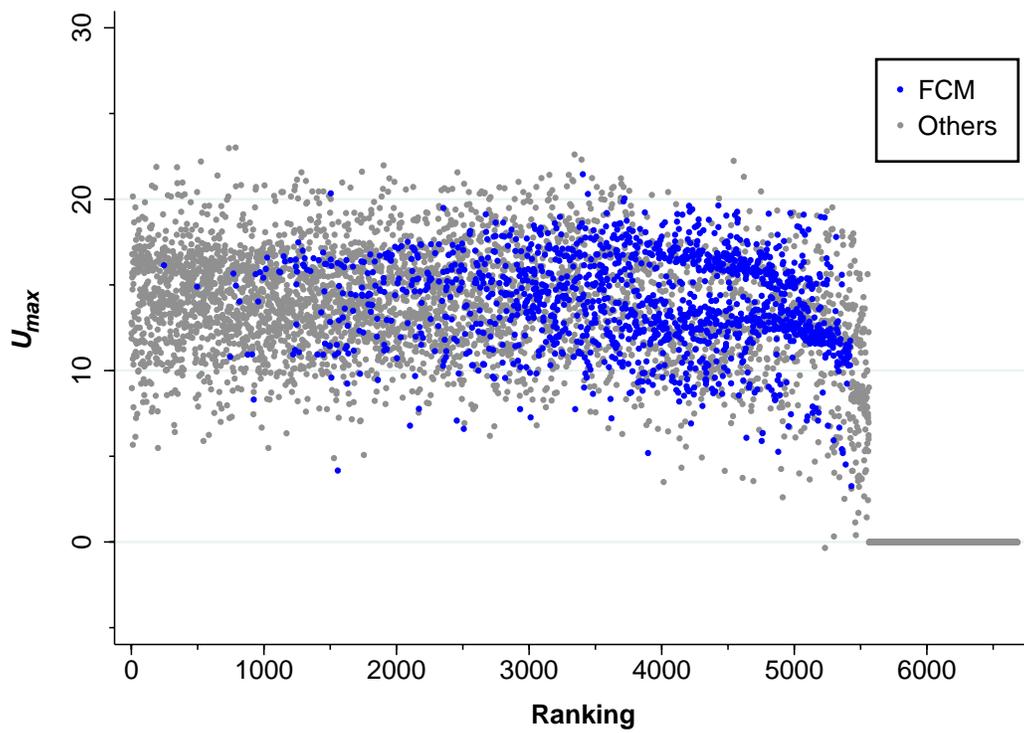


Figure D4. Maximum Attained Utility ( $u_i$ ) versus MIR Ranking ( $R_i$ ) in the Generalized Reduction in Supply Policy Scenario (2B). The blue points represent those candidates who were predicted to choose family and community medicine (FCM).

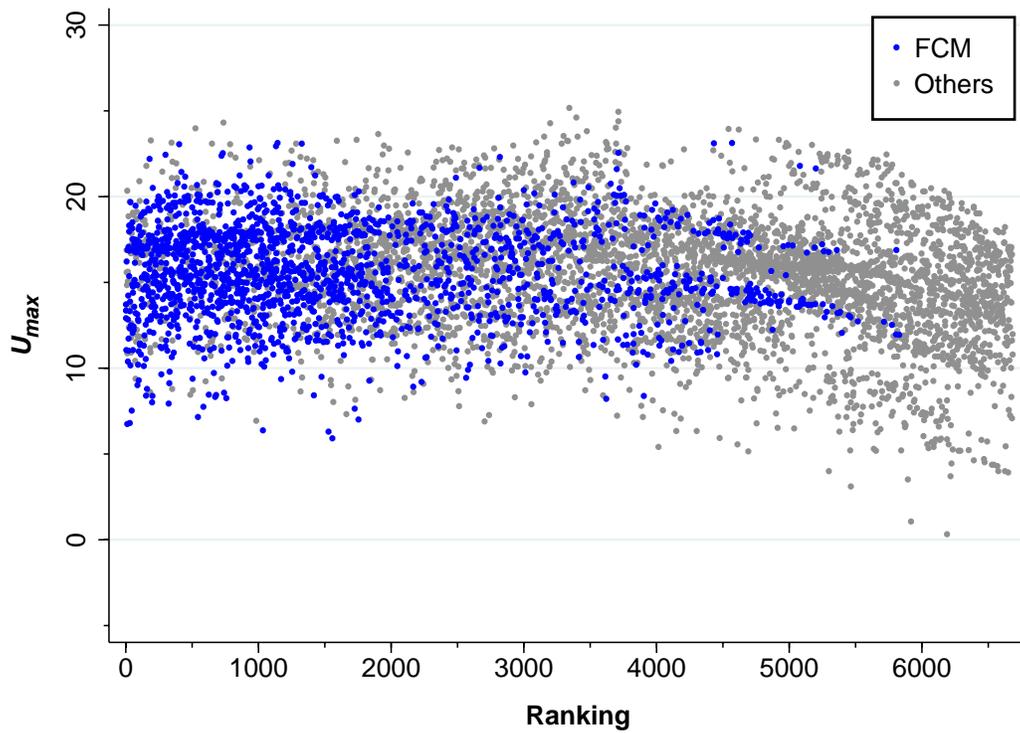


Figure D5. Maximum Attained Utility ( $u_i$ ) versus MIR Ranking ( $R_i$ ) in the Upgrading Family and Community Medicine Policy Scenario (3). The blue points represent those candidates who were predicted to choose family and community medicine (FCM). Prestige among colleagues ( $X_4$ ) was increased from 3.92 to 6.30 on a 10-point scale, while the proportion of income from private practice ( $X_7$ ) was increased from 13.2 to 39.3 on a percentage scale.